DOCUMENT RESUME

ED 055 068	24 SP 005 313
AUTHOR	Merrill, M. David; Tennyson, Robert D.
TITLE	The Effects of Types of Positive and Negative Examples on Learning Concepts in the Classroom. Final Report.
INSTITUTION	Brigham Young Univ., Provo, Utah. Dept. of Instructional Research and Development.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Bureau of Research.
BUREAU NO	BR-0-H-014
PUB DATE	Sep 71
GRANT	OEG-8-70-0006(509)
NOTE	177p.
EDRS PRICE	MF-$0.65 HC-$6.58
DESCRIPTORS	*Concept Teaching; *Educational Research; *Learning Processes; *Research Methodology; *Statistical Analysis

ABSTRACT

Sections 1 and 2 reports on two research studies carried out by the authors, section 3 is a series of abstracts of other research studies stimulated by the authors and section 4 is an appendix containing the instructional materials, tests, raw data, and other items of interest. In the first study the concept "trochaic meter" was taught to 180 college students by means of eight treatment conditions. The independent variables involved presenting a definition or instances or both combined with attribute definition and/or attribute prompting. Dependent variables were errors in correct classification and specified classification. Hypotheses consisted of prediction of particular errors for each treatment and six of the eight hypotheses were supported. In the second study, 100 education psychology students were used in an investigation of four instructional strategies for promoting the acquisition of an infinite concept class. The independent variables were 1) probability level of concept instances; 2) matching of positive instances to a negative instance so that the irrelevant attributes are similar; and 3) divergency of positive instances with one another so that all of their irrelevant attributes differ. The four predicted dependent variables were all statistically significant. (MBM)

ED055068

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

FINAL REPORT

Project No. O-H-014
Contract No. OEG-8-70-0006 (509)

THE EFFECT OF TYPES OF POSITIVE AND NEGATIVE EXAMPLES ON LEARNING CONCEPTS IN THE CLASSROOM

Study 1 Attribute Prompting Variables in Learning Classroom Concepts
Study 2 Concept Acquisition and Specified Errors as a Function of Relationships Between Positive and Negative Instances

M. David Merrill
Brigham Young University
Provo, Utah

Robert D. Tennyson
Florida State University
Tallahassee, Florida

September 1971

This research report herein was performed pursuant to a contract with the Office of Education, U. S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U. S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

SP005 313

Acknowledgements

This project was part of a continuing research program of the Department of Instructional Research and Development at Brigham Young University. Some of the other studies which have been stimulated by this project are reviewed in this report. The findings of these studies and others conducted as part of this research program are now being applied to the development of instructional materials as part of the Instructional Development Program at Brigham Young University. The authors of this report are also preparing *Teaching Concepts: An Instructional Design Guide*, which is a manual for teachers showing them how to apply these findings in the preparation of concept lessons. This manual should be available from a national publisher sometime next year.

The authors wish to express appreciation to the students who served as subjects for these studies and to their instructors who cooperated in the collection of this data. We would also like to thank the extremely competent clerical staff who assisted in the preparation of the instructional materials and the final report.

CONTENTS

Page
Acknowledgements 3
Contents 5
Introduction 7

Study One "Attribute Prompting Variables in Learning Classroom Concepts" 9
  Abstract 9
  Problem 10
  Method 12
  Results 18
  Discussion 20

Study Two "Concept Acquisition and Specified Errors as a Function of Relationships Between Positive and Negative Instances" 30
  Abstract 30
  Problem 31
  Method 35
  Results 40
  Discussion 42

Related Research and Development 53
  "Exemplar and Nonexemplar Variables Which Produce Correct Concept Classification Behavior and Specified Classification Errors" 55
  "Instructional Variables Which Predict Specified Learner Concept Acquisition and Errors" 59
  "Interaction of Anxiety With Performance On Two Levels of Task Difficulty" 65

Supplementary Information

  Appendices for Study One 71
  Appendix A--Introduction Used in All Programs . . 72
  Appendix B--Test Answers 97
  Appendix C--List of Poetry Selections Used in Instance Probability Analysis 99
  Appendix D--Analysis of Variance Tables 111

  Appendices for Study Two 115
  Appendix A--Introduction to all Programs 116
  Appendix B--Introduction to Pretest. 117
  Appendix C--Pretest 118
  Appendix D--Answers to Pretest 124

(Contents Cont)

Appendix E--Introduction to Task 125
Appendix F--Definition of RX_2 Crystals 126
Appendix G--Correct Classification Task 127
Appendix H--Overgeneralization Task 135
Appendix I--Undergeneralization Task 143
Appendix J--Misconception Task 150
Appendix K--Introduction to Nonrelevant Task 158
Appendix L--Nonrelevant Task 159
Appendix M--Introduction to Posttest 163
Appendix N--Posttest 164
Appendix O--Posttest Answers 174
Appendix P--Posttest Probabilities 175
Appendix Q--Posttest Construction 176
Appendix R--Construction of Tasks 177
Appendix S--Conclusion to Program. 178
Appendix T--Analysis of Variance Tables 179

Introduction

This report consists of four sections as indicated in the table of contents.

Sections 1 and 2 consist of reports of the two research studies which were part of the contract. Section 3 is a series of abstracts of other research studies which were stimulated by this contract. These additional studies were completed as dissertations or were funded by Brigham Young University research funds. Section 4 is an appendix containing the instructional materials, tests, raw data, and other items of interest related to this research.

Figures and tables related to a given section are numbered from 1 to N within each section. These figures and tables appear following the text for each section and are printed on yellow paper to facilitate the reader's locating this information.

The contract called for two experimental studies. Both were completed essentially as proposed with some increase in precision of the variables investigated and with a different apparatus than originally proposed.

Departure from proposal. The original proposal called for presenting the instructional materials by means of a computer-controlled slide projector with the student responding by means of a teletype. After expending the funds budgeted for computer rental, and considerable time in an effort to make this apparatus operational, the plan for using this apparatus was abandoned and data was collected using paper and pencil devices. All equipment involved as well as personnel for adapting this equipment were purchased using Brigham Young University funds. Funds from this proposal were used for computer time in an effort to get the system operational.

This change in procedure has little or no effect on the outcome of the studies. The computer would have allowed slightly more control of the stimulus material, but this was not a critical variable in this research. The proposed apparatus would have also facilitated data analysis. This limitation required more time but did not change the amount or type of data analysis employed.

ATTRIBUTE PROMPTING VARIABLES IN LEARNING CLASSROOM CONCEPTS

M. David Merrill
Brigham Young University

Robert D. Tennyson
Florida State University

ABSTRACT

The concept "trochaic meter" was taught to 180 college Ss by means of eight treatment conditions. The independent variables involved presenting a definition (D) or instances (E) or both combined with attribute definition (A) and/or attribute prompting (P). Dependent variables were correct classification and specified classification errors. Hypotheses consisted of prediction of particular errors for each treatment. Six of the eight hypotheses were supported at beyond $p < .01$. The most effective condition for promoting correct classification consisted of D + E + A + P. The last effective condition consisted of D or E alone.

ATTRIBUTE PROMPTING VARIABLES IN LEARNING CLASSROOM CONCEPTS

M. David Merrill
Brigham Young University

Robert D. Tennyson
Florida State University

The need for basic research leading to the development of instructional theory is stressed by Gage (1963), Stolurow (1965), Cronbach (1967), Gagné (1969), as well as other researchers. Each of these investigators has emphasized the special characteristics of such research, its difference from research on the learning process, and its fundamental importance to the development of any systematic instructional technology (courseware).

Gagné's (1970) hierarchical model makes a distinction between eight types of learning outcomes. Category six (concept learning) and category seven (principle learning) were the concerns of this investigation. From an article in Klausmier (1966), Gagné postulates that concept learning and principle learning outcomes require unique forms of instruction. Markle and Tiemann (1969) and Merrill (1971) postulated that adequate concept acquisition (the ability to generalize within a class and discriminate between classes) would result only if exemplars used during instruction differed widely in the irrelevant attributes associated with each; thus promoting generalization within the class. Also, discrimination between classes would result from presenting nonexemplars with irrelevant attributes resembling those associated with given exemplars.

In a study by Tennyson, Woolley and Merrill (1971), independent variables were investigated that predicted concept acquisition and specified classification errors of overgeneralization, undergeneralization, and misconception. The results of their study were based upon three independent variables: probability, matching, and divergency. The group identified as correct classification (concept acquisition) did significantly better on the posttest than the other groups, but an overgeneralization behavior still resulted, i.e., the Ss correctly identified all of the examples as class members, along with identifying some nonexemplars as members of the class. In that condition, where correct classification behavior was hypothesized, some overgeneralization still resulted.

The present investigation is an extension of that design to more adequately control overgeneralization. Gagné's (1970) paradigm for concept learning is the basis for the independent variables manipulated here to increase effectiveness of instruction. His model involves four hierarchical steps: (1) defining the concept, (2) presenting positive and negative instances, (3) identifying attributes of instances, and (4) testing the learner with previously unencountered instances. The Tennyson, etal, (1971) study empirically validated the second step and operationally defined the relationship within positive instances and between positive and negative instances.

Independent Variables

Four independent variables are identified and manipulated in this study: The first variable is definition presentation, which consists of displaying a statement identifying the relevant attributes shared by a set of objects or events in a given class. The second variable is attribute definition presentation, in which each attribute of the concept class is defined and clarified for the learner. An exemplar/nonexemplar presentation is the third variable and consists of displaying exemplars and nonexemplars according to correct classification procedures (see Woolley and Tennyson, 1971; for empirical data, see Tennyson, Woolley and Merrill, 1971). The final independent variable manipulated is a prompting procedure called attribute prompting presentation which consists of explanatory information which indicates class membership and also identifies the attributes for each exemplar or the absence of relevant attributes for each nonexemplar.

Hypotheses

These four independent variables are not manipulated into every possible combination resulting from a completely crossed statistical model. The variables are organized according to logical sets (i.e., can be used in an instructional situation) which are hypothesized to produce certain behavioral outcomes. These outcomes are defined as: correct classification behavior in which S is given previously unencountered exemplars and nonexemplars of a concept,class can correctly identify, undergeneralization behavior, in which S identifies the more obvious exemplars as class members, but indicates that less obvious exemplars are not class members, i.e., he fails to generalize to all members of the class; overgeneralization behavior, in which S correctly identifies all of the exemplars as class members, plus identifying some nonexemplars as members of the class, i.e., the S fails to discriminate between classes; and misconception behavior, in which S falsely assumes that some irrelevant attribute or combination of irrelevant attributes is relevant. The operational consequence is that S fails to recognize exemplars not having this attribute as class members, and indicates that nonexemplars, which do have this attribute, are class members.

The following conditions are hypothesized (Table 1): (1) a definition presentation is hypothesized to produce overgeneralization; (2) a definition presentation plus an attribute definition presentation is hypothesized to produce misconception; (3) an exemplar/nonexemplar presentation is hypothesized to produce overgeneralization; (4) a definition presentation plus exemplar/nonexemplar presentation is hypothesized to produce either correct classification or a slight overgeneralization; (5) a definition presentation with attribute definition presentation plus exemplar/nonexemplar presentation is hypothesized to produce either misconception or overgeneralization (less than hypothesized in condition 2); (6) an exemplar/nonexexemplar presentation plus attribute prompting is hypothesized to produce overgeneralization (less than in condition 3); (7) a definition presentation with exemplar/nonexemplar plus attribute identification is hypothesized to produce correct classification; and (8) a definition presentation with attribute definition presentation with exemplars/nonexemplars presentation plus attribute identification presentation is hypothesized to produce correct classification. This final condition is hypothesized to produce significantly better results than conditions 4 or 7.

Insert Table 1 about here

Method

Learning Task

The instructional objective of the experimental task was: Given a selection of poetry, the S will identify whether it is an example or not an example of trochaic meter. Concept acquisition is required in this task because the S is presented exemplars and nonexemplars in instruction and then required to generalize to previously unencountered exemplars on the posttest, as well as discriminating unencountered nonexemplars. A poetry concept was selected as the task because it is generally used in literature classroom curriculum, and because the irrelevant attributes of poems are infinite. Ninety-five poetry selections were chosen to develop the programs and tests.

To determine which selections to use in the various programs and tests, an instance probability analysis was conducted (Tennyson and

Boutwell, 1971). This procedure has two steps: first, a subjective rating of instances based on difficulty of attributes; and, secondly, an empirical rating based on ability of Ss to identify exemplars from a group of instances containing both positive and negative instances. The subjective analysis involves defining the concept, establishing a list of relevant attributes from the definition and a list of the more common irrelevant attributes (e.g., author, style, period, rhyme, feet, length, etc.), the divergent pairing of exemplars, and matching of exemplars and nonexemplars. An attribute matrix is constructed of the relevant and irrelevant attributes and numerical weights (1 - 5). The poetry selections were given a subjective rating for each attribute and then totaled: the higher the selections' score, the lower the probability. After all of the selections were rated, and a continuum of high to low probability was established, the empirical analysis was conducted.

A sample of Ss were randomly chosen from the target population and presented the definition and poetry selections. The definition used was: Part of the rhythm of a poem is determined by the time between stresses occupied with unstressed syllables or pauses. Denoting the stress patterns is to establish the meter. One of the major meter scansions is named trochee and consists of a stressed syllable followed by an unstressed syllable (marked thusly: /⌣).
Each was asked to study the definition as long as he wished because he could not return to the definition while identifying the selections. Probability was then determined by percentage of Ss correctly identifying each instance. Figure one shows the distribution of the exemplars using a histogram. The probability distribution for the nonexemplars was skewed in favor of correct identification (Figure 2). The correlation between the two analyses was .78.

Insert Figure 1 about here

Insert Figure 2 about here

Procedure

The programs for the 18 treatment conditions followed the same format display: general directions, pretest on poetry (half the Ss), task, and posttest. Ss were administered the program according to individual time arrangements. Upon concluding the general directions, read by the E while S read silently, the S turned to page one and began self-instructional program. In each treatment, half the Ss were given a pretest on poetry which required them to identify examples of trochaic meter. Following the pretest, the Ss received one of the 18 treatments which were previously randomly scrambled. The tasks varied in length, therefore, Ss receiving the shorter programs finished in 15 minutes, while the longer programs required 60 minutes. The posttest was taken immediately following the task. The directions asked the S to read each selection carefully and identify it as an example by writing "yes," or "no" if they thought it was a nonexample.

The programs were printed and stapled together in a self-instructiona booklet. Once the task began, no questions concerning the program were answered by E. Directions required the Ss not to return to previous page and, since the program was nonspeeded, an S could spend as much time per page as desired.

Independent Variables

Manipulation of the four independent variables determines the type of behavior to be elicited from the Ss . A definition, as the first variable, of the concept class based on the relevant attributes, is primary to the task. The definition should be as concise as possible so that prerequisite concepts can be identified. It should be assumed that the Ss can perform the subconcepts identified in the definition (e.g., know what stress marks represent, identify syllables, etc.). Representation of the topic cannot be arbitrary and all examples must be consistent with the identified relevant attributes. The second variable of attribute definition is clarifying of the subconcepts presented in the definition. This variable provides the S with a review, and if necessary, an explanation of the intended meaning of the subconcepts in reference to the definition. In the poetry task used in this experiment, the attribute definition presentation was:

> Words are composed of at least one syllable. Each syllable constitute an elementary sound (diphthong) produced by a single impulse or utteranc and constitutes the word or a part of the word. Adjoining syllables in a word or phrase are marked by abatements, renewals, or reinforcements of the stress so that there is the feeling of separate impulses.
>
> Language patterns have established the pronunciation of words. The syllables are inherent in the language and vary only slightly with varying dialects. Some syllables, such as ing, are generally unstressed or soft. Other syllables, such as pro are generally stressed (given more emphasis when verbalizing). These naturally stressed and unstressed syllables

combine to give variety to language. Poetry carefully uses patterns that can result when words are grouped according to the stressed and unstressed syllables. For example, a line of words that follows this pattern: stress, unstress, stress, unstress; gives a different effect than this pattern: stress, stress, unstress, stress, stress, unstress.

Each of the pattern possibilities have been named for reference. And stressed syllables are marked with a diagonal line above the syllable (e.g., dańc ing). Unstressed syllables have been given a small arc marking (e.g., danc in̆g). The marks are placed directly above the referent syllable. This program will deal with just one pattern used in poetry: Trochee (marked thusly: /◡). It consists of a stressed syllable followed by an unstressed syllable.

Trochaic meter produces a more powerful effect than most other meters. It can convey boredom, frustration, supernatural, or anger, because it starts with a strong beat and it is short. The chants of the American Indians were often set to trochee. The reader must determine, in any poem, which syllables are to be stressed.

The trochaic line often finishes with a strong beat--a masculine ending. Poets generally avoid sustained trochaic measure because of the tendency for it to become monotonous. Children generally enjoy the beat, however, and it is often used in short songs.

The relationship between exemplars and nonexemplars is the third variable investigated. A matched relationship assumes that the exemplar and nonexemplar are to be as similar as possible in their irrelevant attributes. In the case of the poetry task, this would mean having similar rhyme, feet, length, style, author, period, etc. Another relationship assumed in this same variable is between exemplars. Two exemplars presented in correct classification (cf., Tennyson, Woolley, and Merrill, 1971) are divergent when the irrelevant attributes are as different as possible. An example is the following exemplar set with divergent exemplars matched with nonexemplars:

Divergent

Matched

Example:

Out of childhood into manhood
Now had grown my Hiawatha
(Longfellow)

Not an Example:

Come to the crag where the beacon is blazing
Come with the buckler, the lance, and the bow
(Scott)

(Cont.)

(example continued)

Example:	Matched	Pansies, lilies, kingcups, daisies, (Wordsworth)
Not an Example:		Motherly, Fatherly, Sisterly, Brotherly! (unknown)

This variable identifies for the S an example and a nonexample. It does not, however, explain why. The final variable manipulated in this study does describe the rationale for selection. Each instance presented includes a statement which identifies the relevant attributes and why they are relevant. For the nonexamples, the absence of the relevant attribute is noted and explained. An example is given here from the experimental task:

There they are my fifty men and women, Naming me the fifty poems unfinished! (R. Browning)	⟵ Example	By stressing there, the word catches the attention of the reader and leads to the emphasis ending with the exclamation point.
'Tis hard to say if greater want of skill Appear in writing or judging ill. (Pope)	⟵ Not an Example	Words such as 'tis, to, if, -er, and of are not stressed because they do not carry the message as do the stressed syllables.
Boys in sporadic, tenacious droves Come with sticks, as certainly as Autumn. (Eberhart)	⟵ Example	This selection illustrates the masculine line ending wherein the final syllable is stressed.
My mind to me a kingdom is, Such present joys therein I find. (Dyer)	⟵ Not an Example	When reading this poem in a natural manner, the second, fourth, sixth, and eighth syllables are stressed. To alter this produces a strained, vexing effect that was not intended by the poet.

Test

The test was constructed so that the predicted responses of the dependent variables could be analyzed. Thirty selections of poetry were sectioned into three parts with the following format:

1. Convergent high probability exemplar.
2. Convergent low probability exemplar.
3. High probability nonexemplar matched to number 1.
4. Low probability nonexemplar matched to number 2.
5. High probability nonexemplar unmatched.
6. Divergent high probability exemplar paired to number 1.
7. Divergent low probability exemplar paired to number 2.
8. High probability nonexemplar matched to number 6.
9. Low probability nonexemplar matched to number 7.
10. Low probability nonexemplar unmatched.

The thirty selections were randomly scrambled so that no patterns were evident to the Ss. To test the dependent variable of misconception, the grouping of Victorian period poetry was identified as convergent; all other grouping was classified as divergent.

The hypothesized response patterns for each of the dependent variables are given in Table 2. Responses for each S were compared with the predicted score for each dependent variable. S was scored with an error for a given dependent variable when his response to a given item differed from the predicted response. Scores were obtained for the three selections of the test and then added together for the four separate dependent variable conditions. This procedure gave each S four scores; one for each hypothesized dependent variable.

Insert Table 2 about here

Experimental Design

A Solomon Four group experimental design was used so that interaction of pretesting and treatment could be analyzed to control external validity (Campbell and Stanley, 1963). Internal validitv was controlled by random assignment of Ss to the 18 treatments. Since the programs were administered to individual Ss, the basic experimental unit was the S. The n-size for each cell was 10, total N=180. A two-way analysis of variance was used to analyze the data; one main effect was treatment with pretest versus no pretest as the other.

Subjects

The instance probability analysis was conducted with 105 spring semester undergraduate educational psychology *Ss* enrolled at Brigham Young University (BYU). The additional 180 *Ss* who participated in the experiment were randomly chosen from all BYU students enrolled in the spring semester educational psychology classes.

Results

Variable Measures

Four error scores were obtained for each *S* according to hypothesized responses on the dependent variable (Table 1). Table 3 shows the treatment groups, represented by nmemonic labels (see Table 1) and the predicted errors for each dependent variable, i. e., groups DEP and DAEP would make zero errors under the correct classification variable, but it was predicted that groups D, DE, DAE, E, EP would make eight errors, while group DA would make nine errors. No groups were predicted to undergeneralize but the groups were scored on this variable to validate that assumption. Thus, each group was predicted to make significantly fewer errors than the other conditions when its dependent variable was analyzed. Likewise, the other variations in error scores per group were predicted.

Insert Table 3 about here

A two-way analysis of variance was used for each dependent variable. For the correct classification scoring scheme, the main effect of treatments was significant at the .01 level (8, 162 *df*, $F = 23.67$). The pretest versus no pretest effect and interaction were not significant ($p > .05$). With no difference on the pretest, the nine means for the treatment effect were used to determine differences between conditions with the Newman-Keuls Sequential Test. The mean error scores for the treatment groups, according to the dependent variables, are listed in Table 3. The remaining F tests were: Overgeneralization, $F = 15.57$ ($p < .01$); undergeneralization, $F = 29.79$ ($p < .01$) and misconception, $F = 4.83$ ($p < .01$).

Correct Classification

The correct classification dependent variable outcome was hypothesized to result from two instructional procedures. These two conditions (Table 1) were constructed of the definition, exemplar/nonexemplar, and attribute prompting presentations. One group also included an attribute definition presentation. These two conditions did not differ in mean scores ($p > .05$). The two groups did have lower mean error scores ($p < .01$) than the other conditions except for group DE. Only Group DAEP differed from group DE ($p < .05$). Group DE was predicted to overgeneralize with eight errors, while group DEP was hypothesized to make zero errors. The D group, receiving just the definition, and the E group, receiving just the exemplars and nonexemplars, did not differ ($p > .05$). In both cases the mean error scores were slightly better than the control group ($p < .05$). The DAE group varied from the DEP and DAEP groups ($p < .05$), but not from group DE ($p > .05$). Condition group EP, likewise, differed from groups DEP and DAEP ($p < .01$) and group DE ($p < .05$) but failed to vary from group DAE ($p > .05$). In the latter cases, no differences were predicted.

Overgeneralization

The overgeneralization dependent variable was predicted from conditions that did not include sufficient independent variables to teach discrimination. Group D and group E had the lowest mean score on this variable ($p < .01$). Treatment groups DE, DAE, and EP, while different from the previous groups, also, varied from the misconception group DA and the correct classification groups DEP and DAEP ($p < .01$). No differences were predicted between the overgeneralization groups. The control group was significatnly different from the experimental conditions ($p < .01$).

Undergeneralization

None of the treatment condition Ss were predicted to undergeneralize in their responses. The correct classification groups, DEP and DAEP, were more conservative in their responding on the posttest than the other groups ($p < .01$). The only exceptions were groups DE ($p > .05$) and DAE ($p < .05$). Groups D, DA, E, and EP were presented decreased amounts of instruction resulting in increasingly liberal responses. All conditions except the misconception and correct classification were predicted to score the same, yet group D has mean errors higher than all (except group E) experimental conditions ($p < .01$). Only the control group made more responses. The two correct classification treatments (groups DEP and DAEP) were not different from group DE ($p > .05$), even though eight points were predicted. Groups DE and DAE showed no changes ($p > .05$), however, the latter group did differ from the correct classification groups ($p < .05$).

Misconception

The misconception variable was built into the posttest to determine if Ss receiving the definition plus the attribute definition presentation would assume an irrelevant attribute in the attribute definition to be a relevant attribute. There were no differences between the experimental conditions ($p > .05$), the only variance was the control group and the experimental ($p < .01$).

Discussion

Concept acquisition research differs from the traditional concept attainment work in that the former is a deductive instructional system. The learner is given the rule or definition, presented a series of instances and tested with previously unencountered instances to determine transfer. Glass (1968) emphasized this type of research, using school related instructional paradigms as well as real subject matters. This type of investigation would have immediate application to teaching situations. Variables that do effect instruction are basically of three types: stimulus similarity variables, prompting feedback variables, and sequence variables.

In the Tennyson, Woolley and Merrill (1971)study,. the first variable of stimulus similarity was investigated. Their results indicate that a relationship between instances can be operationally defined so as to influence S responses after instruction. Concept attainment research, or other studies in use of negative instances, have not developed the relationship of exemplars and nonexemplars (e.g., Smoke, 1933; Bruner, Goodnow, and Austin, 1956; Donaldson, 1959; Hovland and Weiss, 1953). Tennyson and Merrill (1971) replicated the earlier study of Tennyson etal (1971), but, also, the effect of removing the nonexemplars from the instruction. The results showed that the Ss did not discriminate between exemplars and nonexemplars. The relationship between the two instances forces the learner to focus on the relevant attributes and not be confused by the irrelevant attributes. The results of those two studies (cf. Merrill and Tennyson, 1971) indicated that a definition plus exemplars/nonexemplars in the correct classification format still taught a slight overgeneralization. Divergency, as the other relationship of instances maintains that two exemplars are presented which are as different as possible in their irrelevant attributes This allows the learner to see the range of the concept class.

Probability of instances was an important factor because difficulty of instances was a function of the number and weight of irrelevant attributes. To extend those studies with the addition of the prompting variables was the goal of this investigation.

The prompting variable, by identifying the relevant attribute in each exemplar and the absence of the relevant attributes in nonexemplars, would further explain the definition of the concept. In a study by Woolley (1971), the more information presented to the learner would result in increasingly conservative responses. An anology in statistics is the type II error, in that Ss would be willing to reject a true hypothesis rather than accept a false hypothesis. This study confers this assumption in that the two correct classification groups had the fewest errors on the undergeneralization dependent variables measure. Given a minimum amount of instruction, results in liberal responding closely resembled random choices.

When given just a rule or definition, Group D, the Ss did slightly better than the control group. Adding an explanation of the subconcepts contained in the definition improved performance. Exemplars and nonexemplars without the rule, Group E did the same as the rule-only condition group. The variable which seemed most powerful was the attribute prompting. When this was added to the exemplars/nonexemplars only presentation, the error rate dropped significantly. The definition plus exemplars/nonexemplars condition had the same mean as in the Tennyson, Woolley and Merrill (1971) study ($p > .05$) on the same task. With the addition of the prompting variable and the attribute definition of the subconcepts, the error mean dropped significantly lower.

The results of this study introduce a paradigm of instruction for concept teaching. The variables are not limited to a particular mode of instruction or medium of presentation. The procedures can be readily applied to various subject matter areas or any existing instructional system. Extensions of this study would be on the sequencing variable, such as simultaneous versus sequential presentation of exemplars and nonexemplars; various forms of review, e.g., specific review; adaptive instructional models, which would include the variables investigated for individual differences, e.g., ability, rate, trait, etc.

References

Bruner, J. S., Goodnow, J. and Austin, G. A. *A Study of Thinking.* New York: Wiley, 1956.

Campbell, D. T., and Stanley, J. C. *Experimental and Quasi-experimental Designs for Research.* Chicago: Rand McNally, 1963.

Cronbach, L. How can instruction be adapted to individual differences? In R. Gagné (Ed.) *Learning and Individual Differences.* Columbus, Ohio: Merrill, 1967, 23-39.

Donaldson, J. Positive and negative information in matching problems. *British Journal of Psychology*, 1959, *50*, 253-262.

Gage, N. L. Paradigms for research on teaching. In N. L. Gage (Ed.) *Handbook on Research on Teaching.* Chicago: Rand McNally, 1963.

Gagné, R. M. *The Conditions of Learning* (2nd ed.). New York: Holt, Rinehart, and Winston, 1970.

Gagné, R. M. and Rohwer, W. D. Instructional psychology. *Annual Review of Psychology*, 1969, *20*, 381-410.

Glaser, R. Concept learning and concept teaching. In R. M. Gagné and W. J. Gephart (Eds.) *Learning Research and School Subjects.* Itasca, Illinois: Peacock, 1968.

Hovland, C. I. and Weiss, W. Transmission of information concerning concepts through positive and negative instances. *Journal of Experimental Psychology*, 1953, *45*, 175-182.

Klausmier, H. J., & Harris, C. W. Analysis of concept learning. New York: Academic Press, 1966.

Markle, S. M. and Tiemann, P. W. *Really Understanding Concepts.* Champaign, Illinois: Stipes, 1969.

Merrill, M. D. Necessary psychological conditions for defining instructional outcomes. In M. D. Merrill (Ed.) *Instructional Desigr Readings.* Englewood Cliffs: Prentice Hall, 1971.

Merrill, M. D. and Tennyson, R. D. Concept acquisition as a function of relationships between positive and negative instances. Final Report of USOE Small Contract 0-H-014, September, 1971.

Smoke, K. L. An objective study of concept formation. Psychological Monograph, 1933, 42.

Stolurow, L. M. Model the master teacher or master the teaching model. In J. D. Krumbodtz (Ed.) Learning and the Educational Process. Chicago: Rand McNally, 1965.

Tennyson, R. D. and Merrill, M. D. Correct classification, over-generalization, undergeneralization, and misconception as a function of probability, divergency, and matching of instances in concept acquisition. Submitted Journal of Educational Psychology, 1971.

Tennyson, R. D., Woolley, F. R. and Merrill, M. D. Exemplar and nonexemplar variables which produce correct classification behavior and specified classification errors. Journal of Educational Psychology, 1971, in press.

Woolley, F. R. Effects of the Presence of Concept Definition, Pretraining, Concept Exemplars, and Feedback on the Instruction of Infinite Conjunctive Concepts. Unpublished doctoral dissertation, Brigham Young University, 1971.

Woolley, F. R. and Tennyson, R. D. A conceptual model of classification behavior. Educational Technology, 1971, in press.

Table 1
Hypothesized Outcomes

Treatment	Nmemonic Labels	Dependent Variables
1. Definition Presentation only	D	Overgeneralization
2. Definition Presentation plus Attribute Definition Presentation	DA	Misconception
3. Exemplar/nonexemplar Presentation	E	Overgeneralization
4. Definition Presentation plus Exemplars/nonexemplars Presentation	DE	Overgeneralization
5. Definition Presentation with Attribute Definition Presentation plus Exemplar/nonexemplar Presentation	DAE	Overgeneralization or Misconception
6. Exemplar/nonexemplar presentation plus Attribute Prompting Presentation	EP	Overgeneralization
7. Definition Presentation with Exemplar/nonexemplar plus Attribute Prompting Presentation	DEP	Correct Classification
8. Definition Presentation with Attribute Definition Presentation with Exemplar/nonexemplar Presentation plus Attribute Prompting Presentation	DAEP	Correct Classification

Table 2
Scoring Sheet

	Dependent Variable				
SET # 1	S	M	O	U	C
1. eg #6		+	+	+	+
2. eg #16		+	+	-	+
3. $\overline{eg}$ #3		?	?	-	-
4. $\overline{eg}$ #30		+	+	-	-
5. $\overline{eg}$ #15		?	?	-	-
6. eg #12		-	+	+	+
7. eg #21		-	+	-	+
8. $\overline{eg}$ #8		-	?	-	-
9. $\overline{eg}$ #17		-	+	-	-
10. $\overline{eg}$ #4		-	+	-	-

Note. --Predicted responses according to conditions. M = misconception; O = overgeneralization; U = undergeneralization; C = correct classification; + = S indicates that sentence is a positive instance; - = S indicates this sentence is a negative instance; ? = S could classify as either, no error possible; eg indicates an exemplar $\overline{eg}$ indicates a nonexemplar; # refers to original test item number.

Table 3

Hypothesized Error Responses and Mean Error Scores

	Treatments								Control
Groups	D Over.	DA Misc	DE Over.	DAE Over.	E Over.	EP Over.	DEP Class.	DAEP Class.	
Class.	15.30* 8**	12.55 9	6.40 8	7.27 8	14.90 8	8.75 8	5.25 0	4.36 0	17.80
Over.	4.49 0	11.39 11	9.09 0	8.78 0	5.64 0	8.45 0	12.28 8	13.63 8	15.79
Under.	17.54 14	11.98 9	8.89 14	10.63 14	15.69 14	12.85 14	8.57 6	7.64 6	21.56
Misc.	10.23 9	9.18 0	9.31 11	10.05 11	10.92 11	9.76 11	10.31 9	10.07 9	15.46 9

Note. --The treatment groups are represented by nmemonic labels (see Table 1).
*First rows are the mean error scores. **Second rows are the predicted error scores.

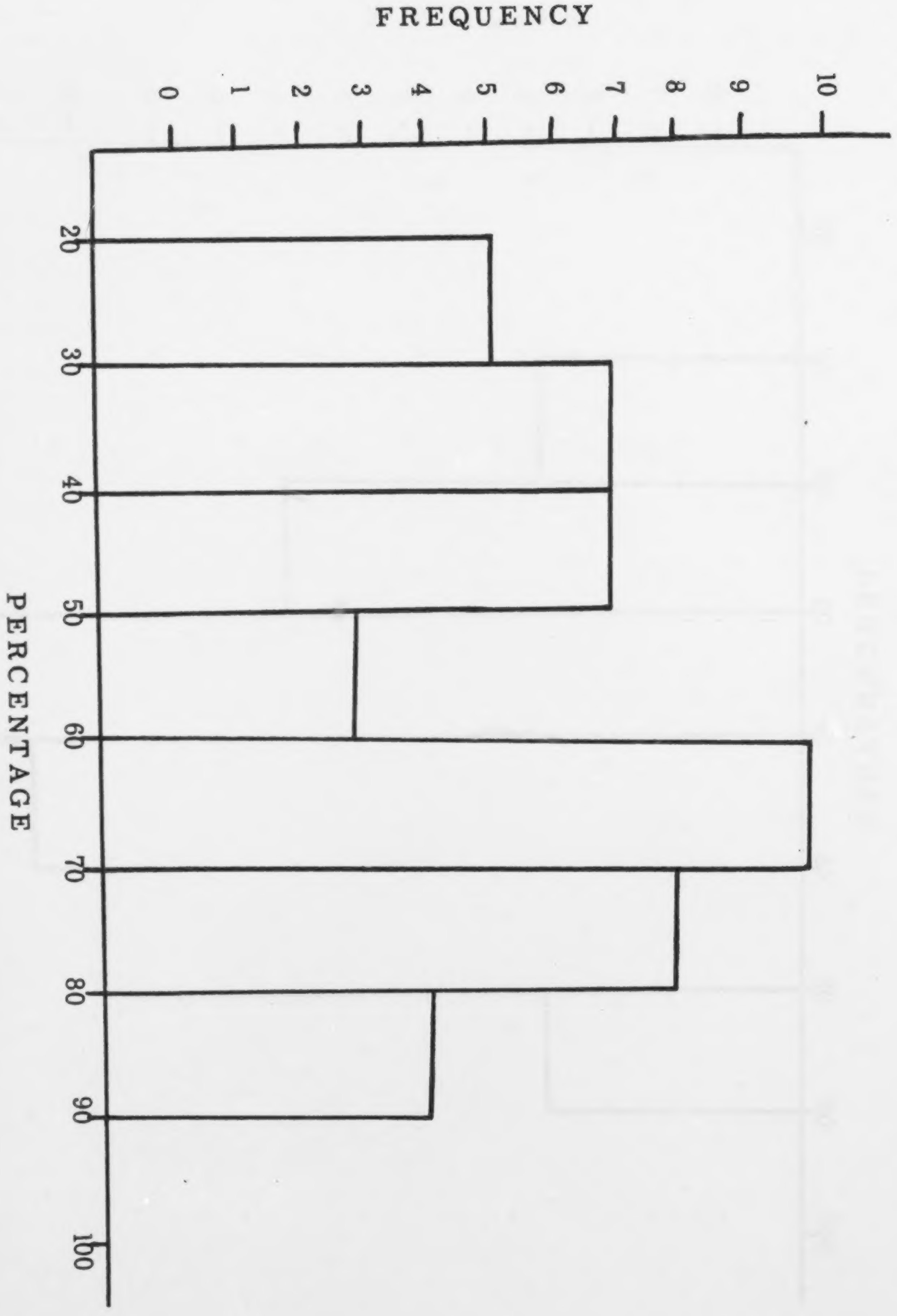


Figure 1. Histogram of correct classification of exemplars by Ss in the Instance Probability Analysis

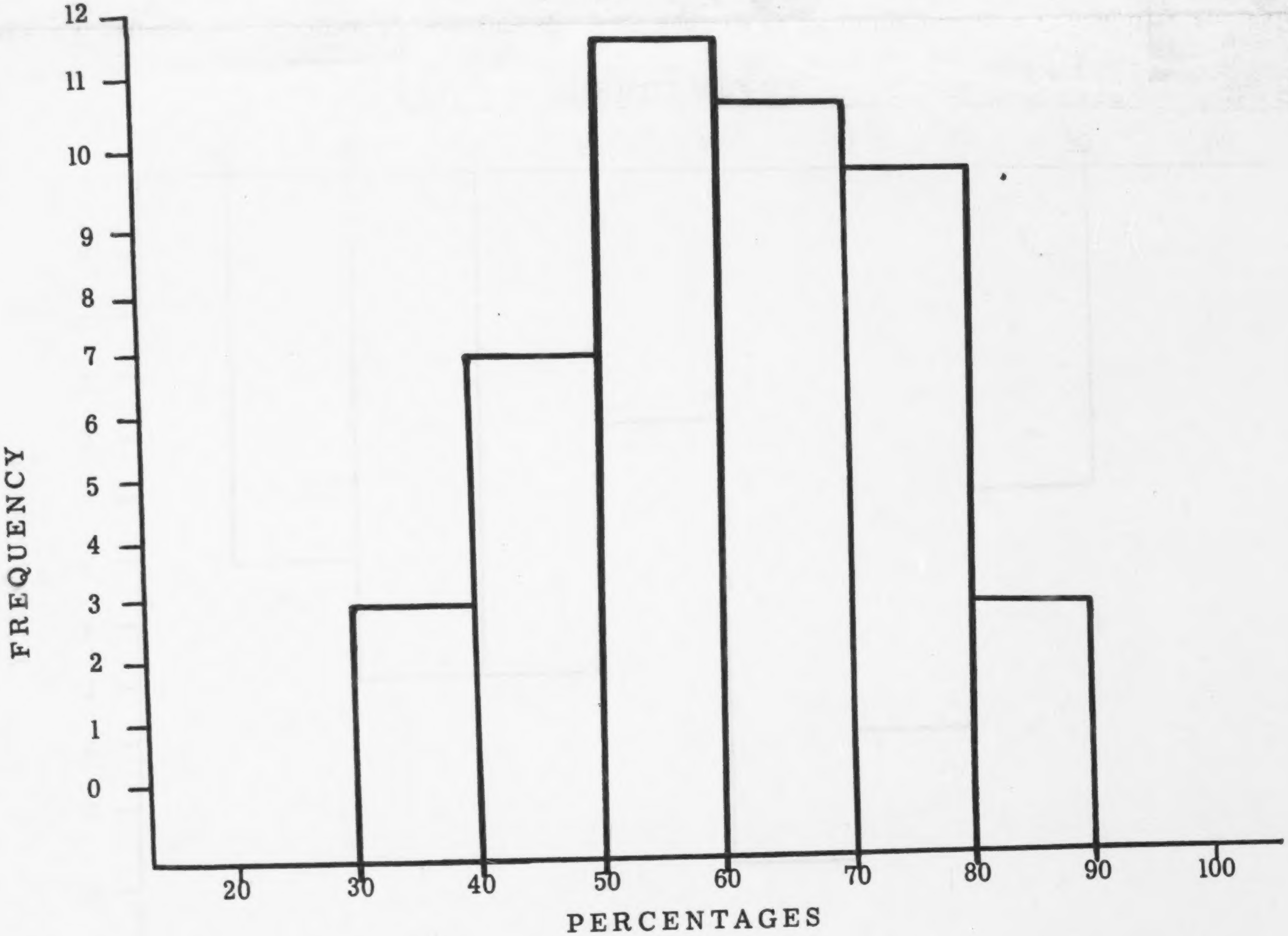


Figure 2. Histogram of correct classification of nonexemplars by Ss in the Instance Probability Analysis.

CONCEPT ACQUISITION AND SPECIFIED ERRORS AS A FUNCTION OF RELATIONSHIPS BETWEEN POSITIVE AND NEGATIVE INSTANCES

M. David Merrill
Brigham Young University

Robert D. Tennyson
Florida State University

ABSTRACT

Four instructional strategies for promoting the acquisition of an infinite concept class were investigated. The independent variables were: 1) *Probability* level of concept instances determined by *Ss* who correctly classify the item as an example or a nonexample; 2) *Matching* of positive instances to a negative instance so that the irrelevant attributes are similar; and 3) *Divergency* of positive instances with one another so that all of their irrelevant attributes differ. Positive instances that share irrelevant attributes are convergent. The manipulation of the independent variables predicted four dependent variables: 1) correct classification; 2) overgeneralization; 3) undergeneralization; 4) misconception. Undergraduate educational psychology students enrolled at Brigham Young University were selected as the 100 *Ss*. The four predicted outcomes were all significant at $p < .01$.

CONCEPT ACQUISITION AND SPECIFIED ERRORS AS A FUNCTION OF RELATIONSHIPS BETWEEN POSITIVE AND NEGATIVE INSTANCES

M. David Merrill
Brigham Young University

Robert D. Tennyson
Florida State University

Controversy has resulted in concept research concerning the value of negative instances (nonexemplars) and their relationship to positive instances (exemplars) in promoting concept attainment. The earliest study dealing with the relationship was Smokes' (1933). He concluded that negative instances were of no value in concept learning. Smoke used an artificial task in which the exemplars and nonexemplars were randomly ordered. This order was changed after each succession through the list. No logical relationship was established between exemplars and nonexemplars. A study which sought to look at the relationship of exemplars was Morrisett and Hovland's (1959) replication of Adams' (1954) study of single vs. multiple task. They found that a variety of positive instances were necessary to effect a transfer (generalization) of concept learning. No attempt was made, however, to establish an operational definition of the relationship between exemplars based on their irrelevant attributes, or on any other criteria. In studies of combined instances, the equivalent attributes of positive and negative instances were found to be poorly utilized by human subjects (Bruner, Goodnow, and Austin, 1956; Donaldson, 1959; Hovland and Weiss, 1953). These studies show the lack of control between exemplars and nonexemplars as does the Smoke study. The concepts were finite with the S attempting to guess the rule (relevant attribute) from a series of instances. Callentine and Warren (1955) studied positive instances and concluded that repetition of one or two instances increased attainment. Luborsky (1945) indicated that eight exposures was more effective than three. These last two studies show that a series of instances is needed, but no mention is made of the difficulty (probability) of the instances or that discrimination of negative instances could be affected by a series which included a combination of exemplars and nonexemplars. Two studies which used negative instances as an integral function of instruction resulted in efficient learning (Huttenlocher, 1962; and Friebergs and Tulving, 1961).

Irrelevant attributes as measures of difficulty have been shown in the studies dealing with ease of attainment of concept classes (Archer,

8S27

Bourne, and Brown, 1955; Brown and Archer, 1957; and Bourne, 1957). Each of these studies found that as the number of irrelevant attributes increased the learning latency and number of errors also increased. They concluded that the number of irrelevant attributes has a linear relationship with difficulty of instances.

Concept acquisition deals with infinite concept classes as contrasted with finite classes as used in concept attainment research (Cronbach, 1967). An infinite class is one in which all of the irrelevant attributes associated with a given exemplar cannot be specified. The procedure for presentation is deductive in that S is told what are relevant attributes and then is given exemplars and nonexemplars prior to the criterion task of identifying class membership. Once an instance has been presented and identified by the S, it is no longer useful as an item to measure this behavior.

Mechner (1965) defined concept acquisition as generalization within a class and discrimination between classes. He pointed out that unless both processes were assessed simultaneously it was not possible to infer concept acquisition. In order to assess concept acquisition, both exemplars and nonexemplars must be presented to the S and his ability to generalize to new exemplars and discriminate them from nonexemplars is observed. Markle and Tiemann (1969) and Merrill (1971) postulated that adequate concept acquisition would result only if exemplars used during instruction differ widely in the irrelevant attributes associated with each; this promotes generalization within the class. Also, discrimination between classes results from presenting exemplars which have irrelevant attributes resembling those associated with given exemplars.

Markle and Tiemann (1969) also postulated that unless the above conditions were met, certain classification behavior errors would result. These are: overgeneralization, undergeneralization, and misconception. Overgeneralization occurs when S correctly identifies all of the exemplars as class members, plus identifying some nonexemplars as members of the class, i. e., the S fails to discriminate between classes. Undergeneralization occurs when S identifies the more obvious exemplars as class members but indicates that less obvious exemplars are not class members, i. e., he fails to generalize to all members of the class. A misconception results when S falsely assumes that some irrelevant attribute or combination of irrelevant attributes is relevant. The operational consequence is that S fails to recognize exemplars not having this attribute as class members and indicates that nonexemplars which do have this attribute are class members.

In a study by Tennyson, Woolley, and Merrill (1971), independent variables were investigated that predicted concept acquisition and specified classification errors. The results of their study were based on three independent variables; probability, matching, and divergency. The probability variable referred to the difficulty of the instances. Probability of each instance was a percentage of students who correctly identified it given only a definition. The rating was taken on a sample from the target population. The matching variable refers to the relationship between exemplars and nonexemplars. A matched condition was defined as instances having similar irrelevant attributes. Divergency referred to the relationship between two exemplars. Exemplars were divergent when their irrelevant attributes were as different as possible. By logically manipulating the three independent variables into four treatment conditions, Tennyson *et al.*, predicted four dependent variables. They were 1) *Correct Classification*, all instances, exemplars and nonexemplars, correctly identified; 2) *Overgeneralization*, nonexemplars similar to class members identified as exemplars; 3) *Undergeneralization*, low probability exemplars identified as nonexemplars; and 4) *Misconception*, exemplars and nonexemplars sharing a common irrelevant attribute identified as class members. The four strategies consisted of presenting to *S* a definition and task according to the hypotheses: 1) IF high to low probability, divergent, and matched, THEN correct classification. 2) IF low probability, divergent, and not matching THEN overgeneralization. 3) IF high probability, divergent, and matching, THEN undergeneralization. 4) IF high to low probability, convergent, and not matching THEN misconception. A score on each independent variable was determined for each *S* on a specially constructed test requiring *S* to identify 30 instances as exemplars or nonexemplars.

In an extension of the above study, Tennyson and Merrill (1971) removed the negative instances from all treatments on a task dealing with the grammatical concept of adverbs. When *Ss* received the task without non-instances, they randomly responded on the posttest. Results indicated that the *Ss* failed to acquire the given concept when presented just positive instances. The failure was both a generalization and discrimination problem.

Independent Variables

The purpose of this study was to operationalize the procedures for selecting the instances. The methodology involves defining the concept class according to relevant attributes. Instances are identified as exemplars if they have all of the relevant attributes. The defining attributes vary in difficulty on two dimensions: the population for whom the instruction is intended; and, difficulty of irrelevant attributes in distracting the S. To account for irrelevant attribute contingencies, the more common ones are identified. The resulting attributes are weighed on a 1-5 scale according to a subjective rating by a subject matter expert. An instance then has an overall probability rating when the weights are totaled: The higher the score, the lower the probability.

Instances are organized into exemplar sets according to the variables of divergency and matching. An exemplar set is composed of two exemplars which differ as much as possible in their irrelevant attributes and two nonexemplars which are matched, one to each exemplar, according to similar or irrelevant attributes. This procedure controls task development by: limiting the definition to the relevant attributes (controlling surplus meaning); it identifies the subconcepts with the assumption that the Ss can perform them; in selecting instances a range of difficulty can be maintained; and the construction of the sets allows a thorough distribution of the concept class. This subjective analysis is then empirically analyzed.

The second component of the instance probability analysis is the collection of the empirical data. A sample of Ss randomly picked from the target population are given the definition of the concept and the list of instances (randomly scrambled) from the subjective analysis. The Ss identify each instance as either an example of the class or not. Probability is then determined by percentage of Ss correctly identifying each instance. If the subjective analysis was rigorous, the instance probabilities should follow a normal curve.

Based on the above methodology for identifying instances, the three independent variables manipulated in this investigation were:

1) Probability: All exemplars and nonexemplars, preceded with a definition of the relevant attributes, were presented to Ss. High probability items are those instances correctly classified by 60% or more of the sample; medium probability are those correctly classified by more than 30% but less than 60%; and low probability are those instances correctly classified by less than 30% of the sample.

2) Matching: An exemplar and nonexemplar are matched when the irrelevant attributes of the two are similar as possible. An unmatched relationship between exemplar and nonexemplar occurs when the irrelevant attributes of the two are as different as possible.

3) Divergency: Two exemplars are divergent when the irrelevant attributes of the exemplars are as different as possible. This relationship assumes the same probability level. A convergent relationship occurs when the irrelevant attributes are as similar as possible.

The four hypotheses resulting from the same manipulation of the three independent variables were also the same (Table 1). Implied in this study are what Stolurow (1969) referred to as "contingency rules," which specify: who is being taught, what is critical, and how the instruction is to be done. Stolurow conceived of contingency rules as "if. . . . and. . . , then. . . . " statements. In the context of this study, the "if" segment contains particular variables identifying the instruction (this is a combination of Stolurow's "if" and "and" segments); the "then" segment contains the predicted outcome or dependent variable.

Insert Table 1 about here

Method

Learning Task

The instructional objective of the task was: Given a picture of a crystal, the S will identify whether or not it is a RX_2 crystal. Concept acquisition was required because the S was presented exemplars and nonexemplars in instruction and then required to generalize to previously unencountered nonexemplars, on the posttest, as well as discriminating unencountered nonexemplars. This concept was chosen for three reasons. First: the concept provides an unlimited number of instances. The molecule arrangements would never appear twice. Second: the concept meets Glaser (1967) and Suppes (1966) concern of using tasks that represent "real" subject-matter learning. Third: Ss would not have previous knowledge of the concept. The definition of the RX_2 crystals was:

Crystals are made up of groups of identical molecules which are comprised of spheres called atoms. The single crystals you are to be tested on may not be complete in and of themselves, but remember that crystals are always symmetrical, so what you don't see may still be present. You must attune yourself to the basic atomic structure or the repeating clusters of atoms. There is a type of crystal called RX_2 which has a two to one ratio in its atomic structure, i.e., for a given atom there will be another two atoms (or cluster of atoms) attached to it in repeating fashion.

Procedure

The programs for the five conditions (the four treatment groups and control group) followed the same format display: general directions, pretest on crystal identification (half the Ss), task, and posttest. Upon concluding the general directions, read by the E while Ss read silently, the Ss turned to page one and began the self-instructional program. Half the Ss in each condition were given a pretest. This required Ss to identify RX_2 crystals from a list containing both positive and negative instances. Following the pretest, a definition and beief explanation of crystals was presented to all Ss except those in the control group. The only instruction for the rest of the program was the presentation of 16 crystal pictures. The format of the exemplar/nonexemplar displays consisted of four sets (two exemplars and two nonexemplars) of crystal pictures--one exemplar and one nonexemplar per page (Figure 1). When finished with the unspeeded program, all Ss took the posttest on RX_2 crystal identification. At the conclusion of the posttest, the S turned in the program and left the testing room.

Insert Figure 1 about here

The programs were printed on high quality paper so that the pictures of the crystals would be clear and free from distortion. This was an important consideration because identification was based on a visual presentation only. Each program was fastened in a colored folder and looked identical on the outside. Responses on the tests were made on I. B. M. answer sheets.

Treatment Programs

The instance probability analysis involved two major components:

1. a subjective rating of instances based on number of irrelevant attributes; and 2. an empirical rating based on ability of random sample of Ss from the target population to identify exemplars from a group of instances. One hundred crystal pictures were selected according to the procedure outlined in the independent variables section. This provided a subjective range of instances which were presented to 100 Ss for the empirical rating. The Ss were randomly divided into four groups and identified the instances while viewing each crystal presented on an overhead projector. The definition was handed each S and was available during the identification. A histogram was constructed to check the distribution of instances.

The four treatment condition programs were developed according to the hypotheses. Programs were administered to several Ss who commented individuality on the program. Several changes were made as a result e.g., the instructions were not totally clear, the definition was shortened by eliminating the attribute definition, and some editorial changes.

For the correct classification program, the exemplar sets were arranged from high to low probability. Page one and two had a high probability set of divergent exemplars with matched nonexemplars followed on page three and four with a high-medium set, page five and sixwith a low-medium set, and page seven and eight with a low probability set. This task was hypothesized to result in a S generalizing to all exemplars on the RX_2 crystals posttest and discriminating the nonexemplars by not identifying them as exemplars.

The overgeneralization program was constructed of only low probability instances. All eight pages of the program were low probability exemplars. However, the nonexemplars from the sets were changed. The new nonexemplars were randomly chosen from other sets. Such an unmatched situation would prevent the S from distinguishing the relevant attribute of RX_2 crystals from some other crystal structure. This situation is the condition found in concept attainment research in which the S sees no relevancy in the negative instances. And as a result, the S fails to develop a discrimination strategy.

The undergeneralization task was constructed of only high probability exemplar sets. Since the nonexemplars were matched, the entire set was used. The first two pages of this program were the same as the correct classification program, with the succeeding pages on an equal level of difficulty. By using only exemplars that had easy subjective ratings and probability ratings above 60%, it was hypothesized that Ss would not generalize to previously unencountered low probability instances because they had not seen the difficult irrelevant attributes.

For the misconception program, the convergent grouping was the irrelevant attributes--solid black molecule with a white mark. This attribute was chosen after discussing with eight Ss what they were looking at as aids in identification. The set included high and low probability ratings. Nonexemplars randomly picked from other sets replaced the matched nonexemplars. With the use of exemplars containing the irrelevant black molecule and with unmatched nonexamples that did not have that particular irrelevant attribute, it was hypothesized that the Ss receiving this condition would respond to the irrelevant attribute whether they were RX_2 crystals or not.

Tests

The pretest consisted of 15 crystals. The items were selected after the four programs and posttest instances were chosen. The posttest items were taken from the same pool of crystals as the programs. For concept acquisition this implies items which were not used in instruction or on the pretest, i.e., previously unencountered instances. Test construction followed this outline:

1. Convergent high probability exemplar.
2. Convergent low probability exemplar.
3. High probability nonexemplar matched to number 1.
4. Low probability nonexemplar matched to number 2.
5. High probability nonexemplar unmatched.
6. Divergent high probability exemplar paired to number 1.
7. Divergent low probability exemplar paired to number 2.
8. High probability nonexemplar matched to number 6.
9. Low probability nonexemplar matched to number 7.
10. Low probability nonexemplar unmatched.

The purpose of the above system is to predict S responses according to type of exemplars and nonexemplars used in instruction. Any number of the above sets can be included in a posttest. For this study three sets were used. The 30 pictures were randomly scrambled so that no patterns were evident to the Ss. To test the dependent variable of misconception, the crystals with black molecules were identified as convergent, all other molecules were classified as divergent. Ss in the misconception treatment condition were hypothesized to classify only convergent high and low probability exemplars and identify as exemplars those matched nonexemplars. The classification treatment group was hypothesized to correctly classify all exemplars on the test. The overgeneralization treatment group was hypothesized to classify not only the exemplars, but also to classify the low probability matched nonexemplars as positive instances. They could also pick high probability matched and unmatched nonexemplars and not be penalized by an error, i.e., a S in this group could have classified all 30 items as exemplars and still follow the predicted results. Undergeneralization Ss were hypothesized to respond only to high probability exemplars.

The hypotheses were stated so that the resulting condition would have fewer errors when its scoring pattern was used against the other groups (Table 3), e.g., the correct classification group would have zero errors, while the overgeneralization group would have eight errors, the undergeneralization group having the most (nine). For each of the other dependent variables, when scored with its pattern, the error would be zero.

The hypothesized response patterns for each of the dependent variables are given in Table 2. S responses were compared with the predicted score for each dependent variable. S was scored with an error for a given dependent variable when his response to a given item differed from the predicted response. Scores were obtained for the three sections of the test and then added together for the four separate dependent variable conditions. This procedure gave the S four scores; one for each hypothesized dependent variable.

Insert Table 2 about here

Experimental Design

A Solomon Four design was used so that interaction of pretesting and the treatment could be analyzed to control external validity (Campbell and Stanley, 1968). This design is constructed so that for each treatment, one half of the S receive a pretest and the other half of the Ss do not. Generalization is increased if the main effect of pretest versus pretest is nonsignificant. A two-way analysis of variance was used with five treatment conditions as one main effect and two levels of pretest as the other. Duncan's New Multiple Range Test was used to determine mean differences among the groups. Internal validity was controlled by random assignment of S to the four programs. Since the programs were administered to individual Ss, the basic experiment unit was S. The n-size for each condition was 10, total N=100.

Subjects

The instance probability analysis was conducted with 100 students enrolled in the undergraduate educational psychology classes at Brigham Young University. The additional 100 Ss in the experimental treatments were students from the same population. Ss were assigned randomly to the five programs. No Ss were dropped from the investigation.

Results

Variable Measures

Four error scores were obtained for each S's responses on the posttest according to the predicted responses on the dependent variables (Table 2). S's responses were compared to the predicted dependent variables and were scored one error for each deviation. Only the correct classification pattern was the correct answer, the other three were based on predicted responses as the results of the manipulation of the three independent variables in the error producing conditions, i.e., overgeneralization, undergeneralization, and misconception. Table 3 shows the treatment groups, represented by capital letters, and the predicted errors for each dependent variable, e.g., the C group would make zero errors with the correct classification variable while O (overgeneralization) group would make eight, the U (undergeneralization) group six, and the M (misconception) group would make nine errors. Thus each group was predicted to make significantly fewer errors than the other three conditions when its dependent variable was analyzed.

Insert Table 3 about here

A two-way analysis of variance for each dependent variable main effects of treatments and pretest versus no pretest, resulted in significance for the first effect ($p < .01$) and nonsignificance for the second ($p > .05$); the interaction was nonsignificant ($p > .05$). With the pretest versus no pretest no difference, four univariate one-way analyses were run with the F=tests (4, 95 df): Classification, $F=4.68$ ($p < .01$); Over generalization, $F=3.39$ ($p < .025$); Undergeneralization, $F=3.30$ ($p < .025$); and Misconception, $F=7.18$ ($p < .01$). The posteriori that used to determine differences among the means for the four separate analyses was Duncan's New Multiple Range Test. The means for the four treatment groups according to the dependent variables are listed in Table 3.

Correct Classification

Concept acquisition was hypothesized to result from an infinite concept class if the exemplars used in instruction were divergently paired in sets with matched nonexemplars. These conditions would provide the S with the behavior to generalize to previously unencountered exemplars and discriminate nonexemplars. Ss receiving the correct classification program on crystals were predicted to identify all the RX_2 crystals on the posttest without responding to other crystals as exemplars, i. e., they would make zero errors on the posttest. The other conditions were hypothesized to make significantly more errors (Table 3). On Duncan's NMRT, the C group made fewer errors than the other groups ($p < 05$). This corresponds to the hypothesis and the predicted responses in Table 3. There were no any differences among groups O, U, and M--slight differences were predicted ($p > .05$). The control group had significantly more errors than the experimental groups ($p < .01$).

Overgeneralization

To promote a discrimination problem for concept acquisition it was hypothesized that unmatched exemplars/nonexemplars would prevent the S from effectively using the negative instances because the irrelevant attributes would have no meaning. The S would not see a matching situation where the irrelevant attributes are the same with only the relevant attributes removed. If presented unmatched nonexemplars, then the S will not discriminate between positive and negative instances when tested with previously unencountered instances. On the crystal posttest it was hypothesized that Ss receiving the overgeneralization treatment would identify more than any of the other groups. The means from the overgeneralization scoring pattern are different (Duncan's NMRT, $p < 05$). The O group did respond to more instances on the posttest which resulted in the fewest number of errors on the overgeneralization scoring pattern. The other experimental groups did not differ ($p > .05$), from each other, with the control group being significantly different ($p < .01$).

Undergeneralization

To produce a condition in concept acquisition where the S can discriminate negative instances, but cannot generalize, it was hypothesized that undergeneralization would occur when using only high probability exemplars, whether they were matched to nonexemplars or not. The S would be instructed with exemplars which had few important irrelevant attributes and when tested with previously unencountered instances which included low probability exemplars, the S would not identify them because of the increasingly complex irrelevant attributes, i. e., the relevant attributes would not be as distinguishable because of the irrelevant attributes. The U group, which received the above treatment, did have fewer errors on the crystal identification posttest than the other groups ($p < .05$). A difference between of 14 points was predicted between the U and O groups. The difference of five points was significant at the .01 level. The C group had fewer predicted errors than the O and M groups, and the mean error scores were different ($p < .05$). The control group had the highest number of errors on the posttest ($p < .05$).

Misconception

Concept acquisition assumes that the S can generalize and discriminate following an instructional situation. In the two error conditions reported above (overgeneralization and undergeneralization), one or the other error was hypothesized to occur. Overgeneralization occurred because of unmatched exemplars/nonexemplars. And undergeneralization resulted from using only high probability exemplars. A combination of these two errors would produce a misconception, i. e., the S would neither generalize to all previously unencountered exemplars or discriminate all unencountered nonexemplars. A misconception error was hypothesized if Ss were presented convergent exemplars, similar irrelevant attributes, and unmatched nonexemplars. A range of probability was included in the hypothesis to account for low probability convergent exemplars. Since the S would receive exemplars with the same irrelevant attributes, S would assume some of these to be relevant. When seeing an exemplar without these irrelevant attributes, S would not identify it. Likewise, upon encountering a nonexemplar with the irrelevant attribute S assumes to be relevant, he would identify it as an exemplar. In this investigation, the irrelevant attribute in which the S was hypothesized to accept as relevant was the black molecule. According to Duncan's NMRT, the M group differed from the other experimental groups ($p < .01$). There were no differences ($p > .05$) among the other conditions, as predicted (Table 3).

Discussion

The relationship of instances in instructional environments seems to be a function of stimulus similarity. This study investigated the role of positive and negative instances using a "real" school subject to establish a paradigm of instruction which has generalizable components. The results indicate that manipulation of the three variables, probability, divergency, and matching, can result in the behavioral errors discussed by Markle and Tiemann (1969). External validity of the variables is strengthened because the results are similar to those of Tennyson, Woolley, and Merrill, (1971) and Tennyson and Merrill (1971).

The independent variable, divergency, dealing with the relationship between exemplars according to their irrelevant attributes was significant. The misconception group was instructed with RX_2 crystals that a pronounced irrelevant attribute which the nonexemplars did not, and as a result identified on the posttest nonexemplars with that same irrelevant attribute as relevant. The other treatment conditions received divergent exemplars and did not respond to the irrelevant attribute when associated with a nonexemplar. Generalization within a concept class would seem to be a function of the divergency variable. By instructing with very different exemplars, Ss transfer more readily when tested with previously unencountered exemplars. The correct classification and overgeneralization groups received a divergent exemplar and responded to the more difficult RX_2 crystal on the posttest. The difficulty of exemplars was determined by the instance probability analysis which subjectively and empirically rated the instances according to ease of recognition.

Probability, as an independent variable, is unique because instances can be rated on difficulty prior to constructing instruction. Individualization of instruction can make sequenencing of easy-to-difficult instances more attuned to the individual by feedback while in an instructional mode. Subjective rating of items has been the usual procedure in all forms of instructional development. The instance probability analysis is a heuristic approach to defining the levels of difficulty of the instances. As irrelevant attributes intensify, the difficulty of the instance increases. Research as early as Bourne's (1957) has shown this linear relationship. Obtaining a subjective analysis on each instance and then constructing matched exemplar pairs with nonexemplars is the first step in deciding which instances to use in the instruction of an infinite concept class. By combining the subjective analysis with a probability rating, the sequencing of exemplar sets eliminates much of the guess work in program development. The most significant difference obtained in this investigation was between the undergeneralization group and the overgeneralization group. An undergeneralization problem was hypothesized when the S received only divergent high probability instances, so the S would not transfer to low probability RX_2 crystals on the posttest. The Ss receiving this treatment made fewer responses on the posttest than any other group. On the contrary, it was hypothesized that the overgeneralization problem, where the S would not discriminate previously unencountered exemplars from nonexemplars, would be promoted by using divergent low probability instances of RX_2 crystals. Ss in this treatment condition responded not only to exemplars but to large numbers of nonexemplars, i. e., they identified more crystals as RX_2 crystals than any other group. The use of the probability variable alone did not cause this problem. Nonexemplars used in the instruction were unmatched with exemplars.

The effect of the matching variable was shown by the increased response to nonexemplars by the overgeneralization group on the posttest. The nonexemplars were unmatched to the exemplars so that Ss failed to recognize the relevant attributes from the irrelevant attributes. When given difficult exemplars, the Ss did respond to the RX_2 crystals on the posttest, i. e., they could generalize to new RX_2 crystal instances, but they could not discriminate from crystals that were not RX_2 crystals. The misconception group had an unmatched relationship between exemplars and nonexemplars, and the Ss failed to distinguish between the relevant and irrelevant attributes of the RX_2 crystals on the posttest. The result of the response patterns on the posttest show that the misconception group responded frequently to crystals having the irrelevant attribute of "black molecules," while not responding to other irrelevant attributes as the overgeneralization group. By using convergent exemplar sets, the Ss focused on a common irrelevant attribute shared by all exemplars and assumed that to be relevant.

Correct classification is a result of the interaction of the three independent variables investigated in this study. There is still some variance which cannot be explained, i. e., the error mean score of the correct classification group, while significantly lower than the other conditions, was still high. Further extensions of this study to account for the variance would include: sequencing of exemplar sets on a more individual basis, i. e., a S might require more of fewer high probability exemplars based upon personality differences (cf., Tennyson and Woolley, 1971); individual probability ratings controlled by adaptive program on a CAI terminal; a more precise measure of the matching variable; and more instructional variables such as defining both relevant and irrelevant attributes.

References

Adams, J. A. Multiple versus single problem training in human problem solving. Journal of Experimental Psychology, 1954, 48, 15-18.

Archer, E.J., Bourne, L.E., Jr., & Brown, F.G. Concept identification as a function of irrelevant information and instructions. Journal of Experimental Psychology, 1955, 52, 316-321.

Bourne, L.E., Jr. Effects of delay of information feedback and task complexity on the identification of concepts. Journal of Experimental Psychology, 1957, 54, 201-207.

Brown, F.G., & Archer, E.J. Concept identification as a function of task complexity and distribution of practice. Journal of Experimental Psychology, 1957, 54,

Bruner, J.S., Goodnow, J. & Austin, G.A. A study of thinking. New York: Wiley, 1956.

Callentine, M.F. & Warren, J.M. Learning sets in human concept formation. Psychology Report, 1955, 1, 363-367.

Campbell, D.T. & Stanley, J.C. Experimental and Quasi-experimental designs for research. Chicago: Rand McNally, 1963.

Cronbach, L.J. How can instruction be adapted to individual differences? In R. Gagńe (Ed.) Learning and individual differences. Columbus, Ohio: Merrill Publishing Co., 1967.

Donaldson, J. Positive and negative information in matching problems. British Journal of Psychology., 1959, 50, 253-262.

Freibergs, V. & Tulving, E. The effect of practice on utilization of information from positive and negative instances in concept identification. Canadian Journal of Psychology, 1961, 15, 101-106.

Glaser, R. Concept learning and concept teaching. In R. Gagńe (Ed.) Learning research and school subjects. Itasco, Illinois: Peacock, 1968.

Hovland, C. I. & Weiss, W. Transmission of information concerning concepts through positive and negative instances. Journal of Experimental Psychology, 1953, 45, 175-182.

Huttenlocher, J. Some effects of negative instances on the formation of simple concepts. Psychological Report, 1962, 11, 35-42.

Luborsky, L. F. Aircraft recognition: The relative efficiency of teaching procedures. Journal of Applicational Psychology, 1945, 29, 385-398.

Markle, S. M. & Tiemann, P. W. Really understanding concepts. Champaign, Illinois: Stipes, 1969.

Mechner, F. Science education and behavioral technology. In R. Glaser (Ed.), Teaching machines and programmed learning, II Data and direction. Washington, D. C.: National Education Association, 1965, 441-507.

Merrill, M. D. Necessary psychological conditions for defining instructional outcomes. In M. D. Merrill (Ed.) Instructional design: readings. Englewood Cliffs, New Jersey, Prentice-Hall, 1971.

Morrisett, L. Jr. & Hovland, C. I. A comparison of three varieties of training in human problem solving. Journal of Experimental Psychology, 1959, 58, 52-55.

Smoke, K. L. An objective study of concept formation. Psychology Monograph, 1933, 42 (4, Whole No. 191).

Stolurow, L. Some factors in the design of systems for computer-assisted instruction. In R. Atkinson & J. Wilson (Eds.) Computer-assisted instruction: A book of readings. New York: Academic Press, 1969, 64-94.

Suppes, P. Mathematical concept formation in children. American Psychologist, 1966, 21, 139-150.

Tennyson, R. D. & Boutwell, R. C. A quality control design and evaluation model for hierarchical sequencing of programmed instruction. NSPI Journal, 1971, 10, 5-10.

Tennyson, R. D. & Merrill, M. D. Correct classification, overgeneralization, undergeneralization and misconception as a function of probability, divergency, and matching of instances in concept acquisition. Submitted Journal of Educational Psychology, 1971.

Tennyson, R. D., Woolley, F. R. & Merrill, M. D. Exemplar and nonexemplar variables which produce correct concept classification behavior and specified classification errors. Journal of Educational Psychology, 1971, in press.

Table 1
Hypotheses Matrix

		Independent Variables Presented			
		Probability	Matching	Divergency	Nonexemplars
Dependent Variable Outcome	1) Correct Classification	All Levels	Matched	Divergent	Yes
	2) Over-generalization	Low Level	Unmatched	Divergent	Yes
	3) Under-generalization	High Level	Matched	Divergent	Yes
	4) Misconception	All Levels	Unmatched	Convergent	Yes

Table 2
Scoring Sheet

Set #1	S	M	O	U	C
1. eg #6		+	+	+	+
2. eg #16		+	+	-	+
3. $\overline{\text{eg}}$ #3		?	?	-	-
4. $\overline{\text{eg}}$ #30		+	+	-	-
5. $\overline{\text{eg}}$ #15		?	?	-	-
6. eg #12		-	+	+	+
7. eg #21		-	+	-	+
8. $\overline{\text{eg}}$ #8		-	?	-	-
9. $\overline{\text{eg}}$ #17		-	+	-	-
10. $\overline{\text{eg}}$ #4		-	+	-	-

Note.--Predicted responses according to conditions. M=misconception; O= overgeneralization; U= undergeneralization; C= correct classification; += S indicates this sentence is a positive instance; - = S indicates this sentence is a negative instance; ? = S could classify as either, no error possible; eg indicates an exemplar; $\overline{\text{eg}}$ indicates a nonexemplar; # refers to original test item number.

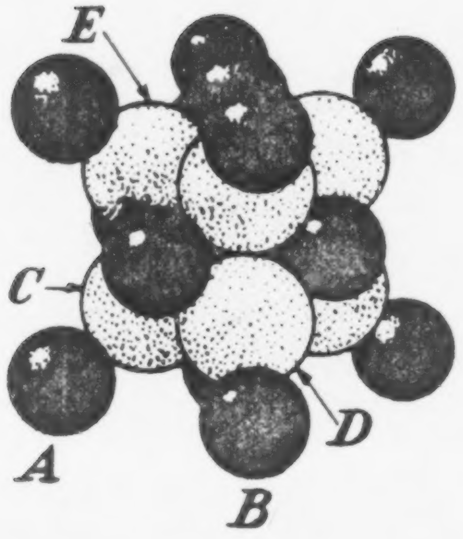
Table 3

Hypothesized Error Responses and Mean Error Scores

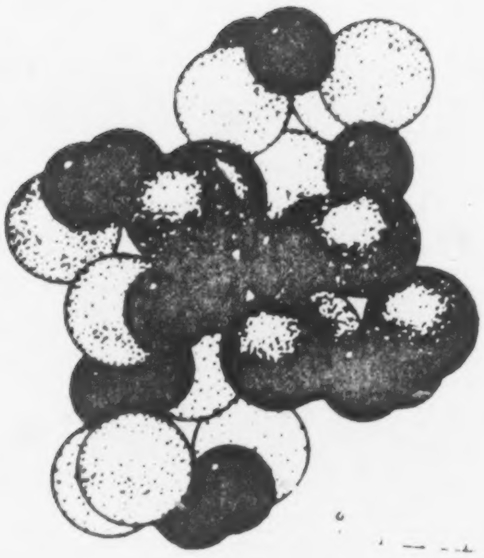
		Treatment Conditions				Control
		C	O	U	M	
Dependent Variables	Class.	10. 55* 0**	13. 25 8	13. 35 6	14. 85 9	18. 10
	Over.	10. 05 3	8. 00 0	10. 05 14	9. 95 11	13. 35
	Under.	12. 75 6	14. 75 14	9. 15 0	14. 75 9	17. 30
	Mis.	10. 65 9	11. 70 11	11. 05 9	7. 10 0	15. 40

Note. --The treatment groups are represented by capital letters: C=Correct Classification; O=Overgeneralization; U=Undergeneralization; and M=Misconception.

*First rows are the mean error scores.
**Second rows are the predicted mean error scores.



EXAMPLE



NOT AN EXAMPLE

Figure 1
The format of exemplar/nonexemplar presentation per page.

Related Research and Development

The present U. S. O. E. grant was profitable in generating extended investigations in concept acquisition. It therefore seems appropriate to include in this final report summaries of the published papers and abstracts of papers submitted to appropriate professional journals. The first study, Tennyson, Woolley and Merrill (1971) was presented at the annual AERA Conference (1971). The reviewer for the *Journal of Educational Psychology* stated, "This paper could prove to be a landmark study in educational psychology." The procedures developed in that study in task construction were used in an aptitude treatment interaction study, Tennyson and Woolley (1971), which will appear in the fall *Journal of Educational Psychology*, and will be presented to the annual APA convention (1971) in division eight, Social and Personality Psychology. A paper summarizing the concept acquisition paradigm, Woolley and Tennyson's (1971), is soon to be published in *Educational Technology*.

The young researchers of the Brigham Young Universities Instructional Research and Development center kept busy with the following studies. Tennyson (1971) investigated the effect of negative instances in instruction of classification behavior. This paper will be presented at the APA convention (1971). The ATI study was followed up by two projects this spring. Boutwell and Tennyson (1971) were investigating anxiety over-time with task difficulties in a group presentation mode. To establish the predictive effect of aptitude and anxiety on task difficulty, Tennyson and Boutwell used a multivariate multiple regression analysis. To define the methodology of the task variable of difficulty, Tennyson and Boutwell (1971) wrote a paper on the Instance Probability Analysis (to be published in *AV Communications Review*).

The Instructional Research & Development center is continuing it's research in the field of what Glaser (1967) calls "school related" topics. The dissemination activity is also going with a home study course by Robert Tennyson on "Concept Learning." A book which will expand this course is being currently written by Tennyson and Merrill.

Additional Papers	Paper
Exemplar and nonexemplar variables which produce correct concept classification behavior and specified classification errors.	9

Merrill/Tennyson	54
Instructional variables which predict specified learner concept acquisition and errors.	
Interaction of anxiety with performance on two levels of task difficulty	10
Conceptual model of classification behavior.	8
Instance probability analysis	20

EXEMPLAR AND NONEXEMPLAR VARIABLES WHICH PRODUCE CORRECT CONCEPT CLASSIFICATION BEHAVIOR AND SPECIFIED CLASSIFICATION ERRORS

Robert D. Tennyson F. Ross Woolley
and M. David Merrill
Brigham Young University

ABSTRACT

Working from a theoretical model, three independent variables were manipulated to produce four predicted dependent variables. The first variable is the relationship between exemplars and nonexemplars. A *matched* relationship exists between exemplars and nonexemplars when the irrelevant attributes are as similar as possible. The difference between the two being the relevant attribute(s). An *unmatched* relationship exists between an exemplar and a nonexemplar when the irrelevant and relevant attributes are different. The second independent variable is the probability rating of exemplars and nonexemplars. High probability exemplars/nonexemplars are those which are correctly classified by a majority of a population when given only a definition identifying the relevant attributes. Low probability exemplars/nonexemplars are those which are not classified correctly by a majority of a population. The third variable is the relationship of exemplars with other exemplars. This relationship is based on the similarity of irrelevant attributes. Two values are used here: *divergent*, the irrelevant attributes are as different as possible; and *convergent*, the irrelevant attributes are as similar as possible.

The three independent variables were combined to predict four dependent variables. The manipulation of the independent variables is based on frequency distribution assumptions and the possible combinations of the independent variables. The dependent variable of *correct classification* is hypothesized to occur when exemplars/nonexemplars are matched, when high and low probability exemplars/nonexemplars are used, and when the exemplars are divergent. *Overgeneralization* is hypothesized to occur when divergent low probability exemplars are unmatched with nonexemplars. *Undergeneralization* is hypothesized to occur when only divergent high probability matched exemplars/nonexemplars are used. And finally, misconception is hypothesized when convergent high and low probability exemplars are unmatched with exemplars.

An empirical task analysis procedure was used to obtain the probability ratings on the exemplars/nonexemplars. The self-instructional programs for each of the four treatment conditions followed one format display. The Ss began the program with a definition of trochaic meter. The format of the exemplar/nonexemplar displays consisted of eight pairs of poetry--two exemplar and two nonexemplar per page. The instances were labeled as exemplar or nonexemplar. Four programs were constructed using the values of the independent variables specified above. The S proceeded to the exam without interruption. The criterion test was constructed so that the predicted responses of the dependent variables could be analyzed, not by mean error scores, but by prearranged predicted response patterns based on each dependent variable (Table 1). S was scored with an error for a given dependent variable when his response to an item differed from the predicted response. This procedure gave the S a score for each of the four hypothesized dependent variables (Table I). Internal validity was controlled by random assignment of Ss to the four programs. The undergraduate BYU student was the basic experimental unit. The n-size for each treatment was 19, total N=76.

Insert Table 1 about here

An analysis of covariance was utilized to test mean significance with S's GPA as the covariate. The means for the four treatment groups are listed in Table 1. There was a significant difference between means for all four treatments, $p < .01$. Correct classification: On the Newman-Keuls Sequential Test and Duncan's New Multiple Range Test at the correct classification (C) group made significantly fewer errors than the other three groups. The other relationships hypothesized were significant either at .01 or .05. Overgeneralization: The Newman-Keuls showed a significant difference at .01 between the overgeneralization (O) group and the undergeneralization (U) group. A difference significant at the .05 existed between O group and misconception (M) and C groups on the Newman-Keuls. Duncan at .01 shows a difference for O with U and M groups, and a .05 existed for O group and C group. Undergeneralization: The multiple comparisons of the undergeneralization error scores are significant at .01 that the U group differed from the O and M groups. There was a .01 between C and U on the correct classification analysis, but here, the difference was only at the .05 on both tests. The other predicted differences are significant in all cases at the .01 level. Misconception: The results followed the predicted variables on all factors at .01 significance. The M group was significantly different from O, U, and G groups.

Precise independent variables were arranged in such a way that predicted dependent variables resulted in all cases. The most significant difference obtained in this study was between the undergeneralization and overgeneralization groups. The undergeneralization group was presented only high probability exemplars and, as a result, responded to few items on the test. On the contrary, the overgeneralization group received only low probability exemplars and responded to practically everything on the test. The independent variable of matched exemplars/nonexemplars can be seen empricially on the increased response to nonexemplars by the overgeneralization and misconception groups. The implication is that discrimination is more effectively taught if the matching of exemplars/nonexemplars is empirically controlled by a task analysis probability rating of both exemplars/nonexemplars. The independent variable of relationship between exemplars with other exemplars according to their irrelevant attributes was significant. The three treatments of classification, undergeneralization, and overgeneralization all received divergent exemplars. Only the misconception group received convergent exemplars.

More work is needed on different subject matter tasks and sample popluations to add external validity to the results gained here. Further research will expand the variables and implications obtained in this study.

TABLE 1

Mean Scores of the Four Dependent Variables

	Means				
	Groups	C	O	U	M
Dependent Variables	Class.	0* 5.685	8* 12.970	6* 9.836	9* 11.980
	Over.	8* 9.014	0* 7.002	14* 11.832	11* 9.255
	Und.	6* 8.553	14* 14.334	0* 6.221	9* 11.627
	Mis.	9* 9.809	11* 10.521	9* 9.759	0* 7.382

*Predicted Errors

INSTRUCTIONAL VARIABLES WHICH PREDICT SPECIFIED LEARNER CONCEPT ACQUISITION AND ERRORS

Robert D. Tennyson

In a study by Tennyson, Woolley, and Merrill (1971) independent variables were found that produced predicted correct classification and specified errors. Mechner (1965) defined concept acquisition as the ability to generalize within a class and discriminate between classes. He maintained that in order to assess concept acquisition, both exemplars and nonexemplars must be presented. In the Tennyson et al. study, correct classification behavior resulted from exemplars which differed widely in the irrelevant attributes, and nonexemplars which had irrelevant attributes resembling those associated with given exemplars. When those conditions were not met, three classification errors resulted: overgeneralization, undergeneralization, and misconception (c. f., Markle and Tiemann, 1969).

Controversy has resulted in concept research concerning the value of negative instances (nonexemplars) and their relationship to positive instances (exemplars) in promoting concept acquisition. Smoke (1933) concluded that negative instances were of no value in concept learning. Morrisett and Hovland (1959), in replication of Adams' (1954) study of single task vs. multiple task, found that a variety of positive instances was necessary to effect a transfer of concept learning. In studies of combined instances, the equivalent attributes of positive and negative instances are found to be poorly utilized by human subjects, (Bruner, Goodnow, and Austin, 1956; Donaldson, 1959; Hovland and Weiss, 1953).

Independent Variables

Based on the theoretical work of Merrill (1971), Markle and Tiemann (1969), Woolley and Tennyson (1971) and the research of Tennyson, Woolley, and Merrill (1971), four independent variables were investigated in this study: 1) Probability: All exemplars and nonexemplars used in instruction are presented to a sample of Ss to determine probability. The Ss receive a definition (relevant attributes) of the concept class prior to classifying the instances. High probability items are those instances correctly identified by 60% or more of the sample. Low probability are those instances correctly classified by less than 40% of the sample. 2) Matching: An exemplar and nonexemplar are matched when the irrelevant attributes of the two are as similar as possible. An unmatched relationship between exemplar and nonexemplar occurs when the irrelevant attributes of the two are as different as possible. 3) Divergency: Two exemplars are divergent when the irrelevant attributes of the exemplars are as different as possible. This relationship assumes the same probability level. A convergent relationship occurs when the irrelevant attributes are as similar as possible. 4) Nonexemplars: The relationship of exemplars/nonexemplars presentation is contrasted with exemplars-only presentation.

Hypothesis

The four variables were combined in various logical groupings to predict four dependent variable outcomes. The independent variables refer to characteristics of exemplars presented to S along with a definition. The predicted S response patterns were measured using additional unencountered exemplars and nonexemplars which the S was asked to classify without confirmation. Two experiments were conducted. Experiment one used nonexemplars in all treatments and the second experiment did not use nonexemplars in all treatments. The hypotheses are summarized in the following statements:

1) If instances represent a range of probability and exemplars are matched to nonexemplars and are divergent with each other, then Ss will correctly classify previously unencountered instances.

1A) If only exemplars are used, the Ss will tend to overgeneralize when classifying previously unencountered instances.

2) If instances are low probability, exemplars are not matched to nonexemplars, and exemplars are divergent with each other, then Ss will tend to overgeneralize when classifying previously unencountered instances.

2A) If only exemplars are used, the Ss will tend to overgeneralize when classifying previously unencountered instances.

3) If instances are high probability, exemplars are matched to nonexemplars, and exemplars are divergent with each other, then Ss will tend to undergeneralize when classifying previously unencountered instances.

4) If instances represent a range of probability, exemplars are not matched to nonexemplars and exemplars are convergent with each other, then Ss will tend to demonstrate a misconception when classifying new unencountered instances.

4A) If only exemplars are used, the Ss will tend to demonstrate a misconception when classifying new unencountered instances.

5) If instances represent a range of probability, exemplars are not matched to nonexemplars, and exemplars are divergent with each other, then Ss will tend to overgeneralize when classifying previously unencountered instances.

5A) If only exemplars are used, the Ss will tend to overgeneralize when classifying previously unencountered instances.

6) If instances are high probability, exemplars are not matched to nonexemplars, and exemplars are divergent with each other, the Ss will tend to undergeneralize when classifying previously unencountered instances.

6A) If only exemplars are used, the Ss will tend to undergeneralize when classifying previously unencountered instances.

Method

Subjects

Ss were seventh grade students from three Utah school districts: Alpine, Provo, and Nebo. Each district provided an alphabetical list of all students from which the Ss were randomly selected.

Task

The grammatical concept of adverbs was used as the task because it is generally used in school English curriculum, and because the irrelevant attributes of this concept are infinite. An instance probability analysis consisting of 120 sentences was conducted to determine probability ratings for each instance.

Procedure

The programs for the twelve treatments (Experiment 1) used the same format display. After reading the general directions,the Ss turned to page 1 and began the self-instructional program. All treatment programs included a pretest on nouns. Half the Ss in each group received in addition a pretest on adverbs. The directions consisted of a short definition of adverbs and method of classifying. The format of the exemplar/nonexemplar displays consisted of eight pairs of sentences--two exemplars and two nonexemplars per page. The instances were labeled as "example" and "not an example." Adverbs were underlined and modification was shown by an arrow.

For experiment two the nonexemplars were removed from the programs used in experiment one. The control group received the pretest and the posttest, the irrelevant task being poetry.

The S proceeded to the posttest without interruption. The criterion test was constructed so that the predicted responses of the dependent variables could be analyzed, not by mean error scores, but by prearranged predicted response patterns based on each dependent variable. S was scored with an error for a given dependent variable when his response to an item differed from the predicted response. This procedure gave the S a score for each of the four hypothesized dependent variables.

Experimental Design

A Solomon Four design was used, i. e., for each treatment half the Ss received a pretest. To control the error variance due to the heterogeneity of the Ss, a covariate pretest was used.

Results

An analysis of covariance was utilized to test mean significance with S's scores on the noun pretest as the covariate. There was a difference between means for all twelve treatments ($p < .01$) in experiment one. No difference resulted between the pretest (adverbs) and unpretest groups ($p > .05$). Correct classification: On the Newman-Keuls Sequential Test and Duncan's New Multiple Range Test, the correct classification (C) group made fewer errors than the other six groups ($p < .01$). The other relationships hypothesized were either $p < .01$ or $p < .05$. Overgeneralization: The Newman-Keuls showed a difference between the overgeneralization(O) group and the undergeneralization (U) group ($p < .01$). Other differences were between O groups and misconception (M) and C groups ($p < .05$). Duncan had similar differences for O group and C group ($p < .05$). Undergeneralization: the multiple comparisons of the undergeneralization error scales showed a difference between U group and O and M groups ($p < .01$). There was s .01 between C and U on the correct classification analysis, but here, the difference was only at the .05 on both tests. The other predicted differences are significant in all cases at the .01 level. Misconception: the results followed the predicted variables on all factors ($p < .01$). The M group was different from O, U, and C groups. In experiment two there was no difference between groups ($p < .05$).

Discussion

Precise independent variables were arranged in such a way that predicted dependent variables resulted in all cases. The most significant difference obtained in this study was between the undergeneralization and the overgeneralization groups. The undergeneralization group was presented only high probability exemplars and, as a result, responded to few items on the test. On the contrary, the overgeneralization group received only low probability exemplars and responded to practically everything on the test. The independent variable of matched exemplars/nonexemplars can be seen empirically by the increased response to nonexemplars by the overgeneralization and misconception groups. The implication is that discrimination is more effectively taught if the matching of exemplars/nonexemplars is empirically controlled by a task analysis probability rating of both exemplars/nonexemplars. The independent variable of relationship between exemplars with other exemplars according to their irrelevant attributes was significant. The three treatments of classification, undergeneralization, and overgeneralization all received divergent exemplars. Only the misconception group received convergent exemplars. The fourth variable of exemplar/nonexemplar presentation is unique since the predicted S response patterns did not result when nonexemplars were excluded. Ss in this treatment condition responded randomly on the posttest, similarly to the control group. When contrasted to the finding in experiment one, the exclusion of nonexemplars in instruction can result in a fourth classification behavior error or random responses.

More work is needed on different subject matter tasks and sample populations to add external validity to the results gained here. Further research will expand the variables and implications obtained in this study.

References

Adams, J. A. Multiple versus single problem training in human problem solving. Journal of Experimental Psychology, 1954, 48, 15-18.

Bruner, J. S., Goodnow, Jacqueline, J., & Austin, G. A. A study of thinking. New York: Wiley, 1956.

Donaldson, M. Positive and negative information in matching problems. British Journal of Psychology, 1959, 50, 253-262.

Hovland, C. I., & Weiss, W. Transmission of information concerning concepts through positive and negative instances. Journal of Experimental Psychology, 1953, 45, 175-182

Markle, S. M., & Tiemann, P. W. Really understanding concepts. Champaign, Illinois: Stipes, 1969.

Mechner, F. Science education and behavioral technology. In R. Glaser (Ed.), Teaching machines and programmed learning, II: Data and directions. Washington, D. C.: National Education Association, 1965, 441-507.

Merrill, M. D. Necessary psychological conditions for defining instructional outcomes. Instructional design: readings. Englewood Cliffs: Prentice-Hall, 1971. In M. D. Merrill (Ed.)

Morrisett, L. Jr. & Hovland, C. I. A comparison of three varieties of training in human problem solving. Journal of Experimental Psychology, 1959, 58, 52-55

Smoke, K. L. An objective of concept formation. Psychological Monograph, 1933, 42 (4, Whole No. 191).

Tennyson, R. D., Woolley, F. R., & Merrill, M. D. Exemplar and nonexemplar variables which produce correct concept classification behavior and specified classification errors. Journal of Educational Psychology, 1971, in press.

Woolley, F. R. & Tennyson, R. D. Conceptual model of classification behavior. Educational Technology, 1971, in press.

INTERACTION OF ANXIETY WITH PERFORMANCE ON TWO LEVELS OF TASK DIFFICULTY

Robert D. Tennyson
Florida State University

F. Ross Woolley
Brigham Young University

Empirical data on the interaction of emotional effects with the cognitive effects of instruction is minimal. In a study by O'Neil, Spielberger, and Hansen (1969) the anxiety state of subjects involved in a self-instructional learning task was significantly related to the difficulty of the subject matter. Their study, with tightly controlled internal validity, was an investigation of Spielberger's (1969) assumption that state anxiety (A-State) differs from trait anxiety (T-State). A-State environmental conditions fluctuate while T-State is the variable related to the more stable anxiety individuals. The study by O'Neil et al. also investigated the Spence (1958)-Taylor (1956) drive theory that high A-State persons would commit more errors on a difficult task than low A-State persons, but this situation would reverse on an easy task. The O'Neil et al. study in part confirmed this assumption when the Ss did increase in anxiety during the difficult task. The study showed an error reverse interaction within the difficult task.

Independent Variable

This study investigated the independent variable of task difficulty in connection with the interaction of individual anxiety states and instructional treatment. The independent variable included two levels of difficulty, an easy and a difficult task. An easy task was defined as one in which the displayed exemplars (positive instances) and nonexemplars (negative instances) were of high probability--a majority of a given population when given the definition (list of relevant attributes) can classify the previously unencountered items as members of a given class. The difficult task was defined as one composed of low probability exemplars and nonexemplars--defined as those previously unencountered instances which a majority of a given population when given the definition cannot classify correctly.

Hypotheses

The dependent variable was the S's error rate recorded for each level of the S's A-State. Ss with a low measured A-State were hypothesized to make fewer errors on the difficult task than Ss with a high measured A-State. Conversely, high A-State Ss were hypothesized to make fewer errors on the easy task than low A-State Ss. Thirdly, S's measured A-State score was hypothesized to increase following the difficult task and to decrease with the easy task. Thus, the study hypothesizes that a disordinal interaction exists between task difficulty and the S state anxiety.

Method

Subjects

The instance probability analysis was conducted with 35 students enrolled in an undergraduate educational psychology class at Brigham Young University (BYU). The program sessions involved 29 randomly selected BYU undergraduate general psychology students (12 males and 17 females).

Task

The self-instructional program for this study was concept acquisition. The concept's relevant attribute was poetic trochaic meter. An instance probability analysis involving 85 pieces of poetry was conducted to determine correct classification probability ratings for each piece of poetry. The Ss studied a definition of trochaic meter and then identified each poetic selection as either an example or not an example. A frequency distribution of correct classifications resulted which resembled a normal one-tailed curve. High probability selections were those correctly defined by the Ss at a frequency of 60% level and above; low probability selections were identified at 40% level for exemplars and at 50% level for nonexemplars. The task was composed of two parts: a difficult section of poetry selections, and an easy section of poetry selections.

Anxiety Measures

The Spielberger, Gorsuch and Lushene (1969) State-Trait Anxiety Inventory (STAI) was used to measure A-Trait and A-State. The A-Trait (Form X-2) Scale asks Ss to indicate how they "generally feel", while the A-State (Form X-1) Scale requires Ss to indicate how they feel "at this moment." Systolic blood pressure (SBP) was taken using a blood pressure cuff (sphygomomanometer) and stethescope.

Experimental Design

This study utilized a multiple treatment experimental design (Campbell and Stanley, 1963). Three Es worked with a maximum of eight Ss per session. A session consisted of four periods in which the SBP and STAI A-State Scale were administered.

Procedure

During the pretask period each S took the STAI A-Trait Scale followed by the STAI A-State Scale and the SBP. The S was then given directions and the definition. The second period was the difficult task followed by the SBP and the STAI A-State Scale. In the third period the S again received general directions and proceeded with the easy task followed by the SBP and the STAI A-State Scale. During the fourth period the S was alone for three minutes. The SBP was taken and the STAI A-State Scale was administered a fourth time.

Results

The dependent variable measures were STAI A-State Scale and A-Trait Scale scores, SBP readings, and task errors. The STAI A-Trait Scale score was correlated .44 with the second (difficult period) STAI A-State Scale score. An r of .62 was calculated between STAI A-Trait Scale and the third measure (easy period) of STAI A-State Scale score. Ss were separated into low and high anxiety by the median score of 39 on the first STAI A-State Scale. There was a correlation of .92 between the Ss initial classification on anxiety level and his classification on second (difficult) and third (easy) STAI A-State Scales.

An analysis of variance of repeated measures was used with a Newman-Keuls Sequential Test to obtain significant differences of STAI A-State Scale means. There was a difference between anxiety periods ($F=6.58$, $df=3/84$, $p<.01$). The STAI A-State Scale scores increased from the pretask period to the difficult task and decreased following the easy task with no change within the posttask period. The mean differences were between the difficult period and the other period ($p<.01$), but there were no differences between the pretask and easy task or posttask ($p>.05$). Blood pressure readings did not vary from period to period ($F=1.07$, $df=3/84$, $p>.05$).

A two-factor design was used to analyze the interaction. Task treatment, at two levels, was administered to all Ss, while the anxiety effect was controlled with two different groups. For the main effects, there were no differences ($p>.05$) for either the Between-Ss on anxiety or the Within-Ss on task difficulty; however, the interaction was ($F=11.26$, $df=1, 28$, $p<.01$). The error means were plotted according to anxiety groups and task difficulty (Figure 1). To test for differences between the means withing the anxiety level, the Least Significant Difference (LSD) test was used because of its decreased probability of making a Type II error. The low A-State group differed between tasks ($p < .05$). A difference was found between the mean errors of the tasks for the high A-State group ($p < .05$). The low A-State group made fewer errors on the difficult task than the high A-State group ($p < .05$). The disordinal interaction was complete with the high A-State group making fewer errors on the easy task than the low A-State group ($p < .05$). To check the results of the LSD test, an Individual Degree of Freedom test was included; the results were similar ($p < .05$).

Insert Figure 1 about here

Discussion

The disordinal interaction of measured high anxiety individuals with low anxiety individuals on the two levels of task difficulty follows the assumptions of the Spence-Taylor drive theory. Since the data did result in a significant disordinal interaction, several implications for instruction are evident. Individuals who do have anxiety increases during difficult tasks might be expected to perform more efficiently if they receive instruction geared for slower increases of difficulty. The opposite expectation would be indicated for low state anxiety individuals. For these individuals, to be instructed with an easy task would produce a less effective means of learning. The disordinal interaction of state anxietv and instructional complexity cannot be ignored in instructional systems.

References

Campbell, D. T. & Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago: Rand-McNally, 1963.

O'Neil, H. F., Spielberger, C. D. & Hansen, D. H. Effects of state anxiety and task difficulty on computer-assisted learning. Journal of Educational Psychology, 1969, 60, 343-350.

Spence, K. W. A theory of emotionally based drive (D) and its relation to performance in simple learning situations. American Psychologist, 1958, 13, 131-141.

Spielberger, C. D. Theory and research on anxiety. In C. D. Spielberger (Ed.), Anxiety and behavior. New York: Academic Press, 1966.

Spielberger, C. D. Gorsuch, R. & Lushene, R. The state-trait anxiety inventory, Test Manual for Form X. Palo Alto, California: Consulting Psychologists Press, 1969.

Taylor, J.A. Drive theory and manifest anxiety. Psychological Bulletin, 1956, 53, 303-320.

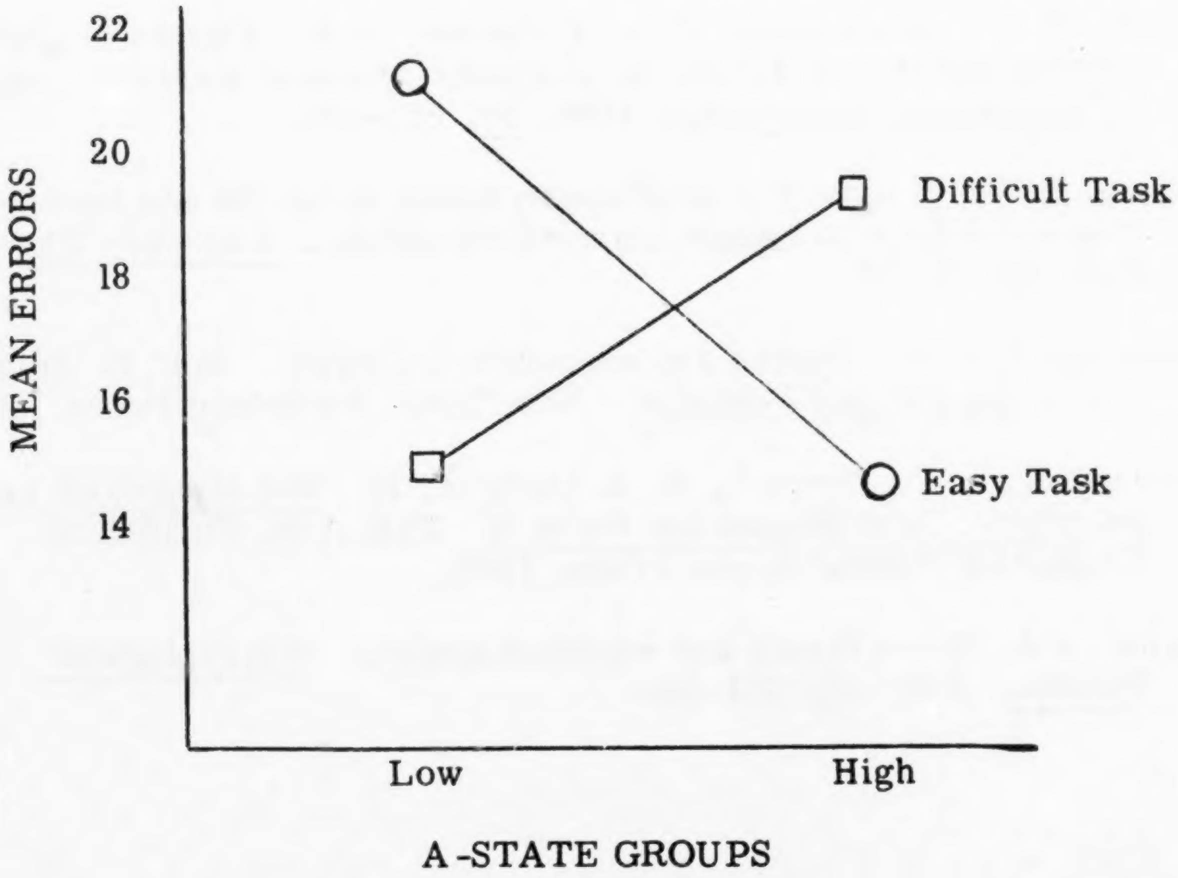


Fig. 1. Interaction of STAI A-State Scale group and mean errors on task period.

List of Appendices for Study One

Appendix Page

A Introduction Used in All Programs

A-1 Pretest Used in Groups 1, 3, 5,..., 17
A-2 Introduction to Trochaic Meter Program Used in All Groups Except 9, 10, 11, 12 17, 18.
A-3 Definition Presentation Used in All Groups Except 9, 10, 11, 12, 17,18.
A-4 Attribute Definition Used in Groups 3, 4, 7, 8, 15, 16.
A-5 Exemplars-Nonexemplars Presentation Used in Groups 5, 6, 7, 8, 9, 10, 11, 12.
A-6 Exemplars-Nonexemplars Plus Attribute Identification Used in Groups 11, 12, 13, 14, 15, 16.
A-7 Nonrelevant Task Used in Groups 17, 18,
A-8 Posttest Used in All Groups
A-9 Introduction to Groups 9, 10, 11, 12

B Test Answers

B-1 Answers to Pretest
B-2 Answers to Posttest

C List of Poetry Selections Used in Instance Probability Analysis

C-1 Instance Probability Analysis for Exemplars
C-2 Instance Probability Analysis for Nonexemplars

D Analysis of Variance Tables

D-1 Correction Classification
D-2 Overgeneralization
D-3 Undergeneralization
D-4 Misconception

Appendix A Introduction Used In All Programs

Dear Student:

You have been randomly selected from all educational psychology students to participate in a research grant funded by the United States Office of Education. The instructions you are to receive have been developed to attain specific information. Please follow all directions carefully.

This program is designed to be self-instructional. The experimenter cannot answer any questions. If for some reason you cannot continue with the program, please take the program to the experimenter and leave quietly. There are no timed breaks, so once you begin, continue until finished.

When you have finished reading the above, please continue to the next page.

Appendix A-1 Pretest Used in Groups 1, 3, 5, . . . , 17

The next few pages will contain selections of poetry which you are to identify as either being examples of trochaic meter or not an example of trochaic meter. This is a pretest and it is used to determine your knowledge of trochaic meter before you are given instructions. We assume that most students are not familiar with trochaic meter.

After reading each selection, you are to respond by writing "yes" opposite the correct number if an example, or "no" if not an example.

When you have finished reading the above, please continue to the next page.

PRETEST

1. Now the day is over
Night is drawing nigh
Shadows of the evening
Steal across the sky.
(Sabine Baring-Gould)

2. Take thought:
I have weathered the storm
I have beaten out my exile.
(Pound)

3. Through the noises of the night
She floated down to Camelot:
(Tennyson)

4. Sure solacer of human cares,
And sweeter hope, when hope despairs!
(Bronte)

5. From ghoulies and ghosties
Go out ghastly grumblings.
(Anonymous)

6. And we just made it out of the Big Muddy
With the captain dead and gone.
(Seeger)

7. "Spanish ships of war at sea! We have
sighted fifty-three!"
(Tennyson)

8. Says he, "Dear James, to murder me
Were a foolish thing to do,
For don't you see that you can't cook me,
While I can--and will--cook you!"
(Gilbert)

9. I don't know why she didn't like my saying
that. She gave me her plaintive smile and her
beautiful eyes filled with tears.
(Maugham)

After you have answered all of the above questions, continue to the next page.

10. Glory be to God for dapple things--
(Hopkins)

11. Yet, as if grieving to efface
All vestige of the human race,
On that lone shore loud moans the sea,
But non, alas! shall mourn for me!
(Wilde)

12. The God of love my Shepherd is,
And He that doth me feed,
While He is mine, and I am His,
What can I want or need?
(Herbert)

13. Have I not passed thee on the wooden bridge
Wrapt in thy cloak and battling with the snow.
They face toward Hinksey and its wintery ridge?
(unknown)

14. Could I but live again
Twice my life over,
Would I not strive again?
(Browning)

15. He says we are beggars.
(Randall)

16. Oh, to be in England
Now that April's there.
(Browning)

17. Wild Spirit, which art moving everywhere;
Destroyer and preserver; hear, O, hear!
(Shelley)

18. When they shot Malcolm Little down
On the stage of the Audubon Ballroom,
When his life ran out through bullet holes
(Like the people running out when the murder began)
His blood soaked the floor.
(Patterson)

After you have answered all of the above questions, continue to the next page.

19. A crack was in the glass
(unknown)

20. Since I can never see your face,
And never shake you by the hand,
I send my soul through times and space
To greet you. You will understand.
(Flecker)

21. The readers of the *Boston Evening Transcript*
Sway in the wind like a field of ripe corn.
(Eliot)

22. Take her up tenderly,
Lift her with care;
(Hood)

23. But the wind can so easily whip
  The still water to foam,
And the harmless bay waves turn
  To storming gray boulders that pound.
(unknown)

24. Memory tells me of the many times
We were three in a room, trapped
By the all-constricting walls of a
summer shower.
(Lee)

25. He says we are beggars and I say
When it comes to love we are orphans
We are all misplaced or displaced
Persons from another war.
(Randall)

26. Tasks in hours of insight willed
Can be through hours of gloom fulfilled.
(Arnold)

27. Soft and easy is thy cradle
Coarse and hard thy Savior lay.
(unknown)

After you have answered the above questions continue to the next page.

28. My mind to me a kingdom is,
Such present joys therein I find.
(Dyer)

29. When the breakers are roaring like beasts on the floor
of the Bourse,
And the poor have the sufferings to which they are
fairly accustomed,
And each in the cell of himself is almost convinced
of his freedom;
A few thousand will think of this day.
(Auden)

30. The wine of life keeps oozing drop by drop.
(Fitzgerald)

After you have answered all of the above questions, continue to the next page.

Appendix A-2 Introduction to Trochaic Meter Program Used in all Groups Except 9, 10, 11, 12, 17, 18.

The following pages will present to you a program to teach the concept of trochaic meter. The programs vary in length according to a random selection. Study each program carefully.

When you have finished reading the above, please continue to the next page.

Appendix A-3 Definition Presentation Used In All Groups Except 9, 10, 11, 12, 17, 18.

Part of the rhythm of a poem is determined by the time between stresses being occupied with unstressed syllables or pauses. Denoting the stress patterns is to establish the meter. One of the major meter scansions is named trochee and consists of a stressed syllable followed by an unstressed syllable (marked thusly:/◡).

When you have finished studying the above, please continue to the next page.

Appendix A-4 Attribute Definition Used In Groups 3, 4, 7, 8, 15, 16.

Words are composed of at least one syllable. Each syllable constitutes an elementary sound (diphthong) produced by a single impulse of utterance and constitutes the word or a part of the word. Adjoining syllables in a word or phrase are marked by abatements, renewals, or reinforcements of the stress so that there is the feeling of separate impulses.

Language patterns have established the pronunciation of words. The syllables are inherent in the language and vary only slightly with varying dialects. Some syllables, such as ing, are generally unstressed or soft. Other syllables, such as pro, are generally stressed (given more emphasis when verbalizing). These naturally stress, unstress syllables combine to give variety to language. Poetry carefully uses patterns that can result when words are grouped according to the stressed and unstressed syllables. For example, a line of words that follows this pattern: stress, unstress, stress, unstress; gives a different effect than this pattern: stress, stress, unstress, stress, stress, unstress.

Each of the pattern possibilities have been named for reference. And stressed syllables are marked with a diagonal line above the syllable (eg. danć ing). Unstressed syllables have been given a small arc marking (eg. danc ĭng). The marks are placed directly above the referent syllable. This program will deal with just one pattern used in poetry: Trochee (marked thusly: /◡). It consists of a stressed syllable followed by an unstressed syllable.

Trochaic meter produces a more powerful effect than most other meters. It can convey boredom, frustration, supernatural, or anger because it starts with a strong beat and it is short. The chants of the American Indians were often set to trochee. The reader must determine, in any poem, which syllables are to be stressed.

The trochaic line often finishes with a strong beat--a masculine ending. Poets generally avoid sustained trochaic measure because of the tendency for it to become monotonous. Children generally enjoy the beat, however, and it is often used in short songs.

When you have finished studying the above, please continue to the next page.

Appendix A-5 Exemplars-Nonexemplars Presentation Used in Groups 5, 6, 7, 8, 9, 10 11, 12.

Example:

Out of childhood into manhood
Now had grown my Hiawatha.
(Longfellow)

Not an Example:

Come to the crag where the beacon is blazing,
Come with the buckler, the lance, and the bow.
(Scott)

Example:

Pansies, lilies, kingcups, daisies.
(Wordsworth)

Not an Example:

Motherly, Fatherly, Sisterly, Brotherly!
(unknown)

Example:

Lay a garland on my hearse of the dismal dew
Maidens, willow branches bear, say I died true.
My love was false, but I was firm from my hour of birth;
Upon my buried body lay lightly, gently, earth.
(Fletcher)

Not an Example:

Nor marble, nor the gilded monuments
of princes, shall outlive this powerful rime;
But you shall shine more bright in these contents
Than unswept stone, besmeared with sluttish time.
(Shakespeare)

Example:

We stripped and dived and found his body
Stuck in the old quicksand.
(Seeger)

When you have finished studying the above, please continue to the next page

Not an Example:	And we just made it out of the Big Muddy With the captain dead and gone. (Seeger)
Example:	There are they, my fifty men and women Naming me the fifty poems unfinished! (R. Browning)
Not an Example:	'Tis hard to say if greater want of skill Appear in writing or in judging ill. (Pope)
Example:	Boys in sporadic, tenacious droves Come with sticks, as certainly as Autumn. (Eberhart)
Not an Example:	My mind to me a kingdom is, Such present joys therein I find. (Dyer)
Example:	I want to know what is really going on. (Resendez)
Not an Example:	Alone, alas, He sat. (unknown)
Example:	I will go up to the mountain And there I will light a fire. (Austin)
Not an Example:	My mind to me a kingdom is Such present joys therein I find. (Dyer)

When you have finished studying the above, please continue to the next page.

Out of childhood into manhood Now had grown my Hiawatha. (Longfellow)	Example	Trochaic because the marked syllables are the most important. It carries the Indian chant quality.
Come to the crag where the beacon is blazing, Come with the buckler, the lance, and the bow. (Scott)	Not an Example	Because the two unaccented syllables follow the stressed beat, the line is not trochaic.
Pansies, lilies, kingcups, daisies. (Wordsworth)	Example	Speaking in normal language, we automatically make these words trochaic by slightly emphasizing the first syllables.
Motherly, Fatherly, Sisterly, Brotherly! (unknown)	Not an Example	Each of these words has three syllables and only the first is accented.

When you have finished studying the above,
please continue to the next page.

Appendix A-6 Exemplars-Nonexemplars Plus Attribute Identification Used In Groups 11, 12, 13, 14, 15, 16.

I want to know what is really going on. (Resendez)	Example	The modern "protest poets" are using the trochaic emphasis to indicate their strong feeling. The stressed beats are exaggerated more in these lines than in many other poems.
Alone, alas, He sat. (unknown)	Not an Example	The short a is seldom stressed as a long a might be: (a′corn). The structure of the words and natural English inflections indicate that the beat of this example is not trochaic.
I will go up to the mountain And there I will light a fire. (Austin)	Example	The last phrase gives the reader the main key to the meter of this selection. I* Light* Fire forms the basic message and each is necessarily stressed.
My mind to me a kingdom is Such present joys therein I find. (Dyer)	Not an Example	This selection is not trochaic because the words MIND and ME need to be stressed over MY and TO. The word value must be carefully considered when deciding the meter of a poem.

When you have finished studying the above, please continue to the next page.

There they are my fifty men and women, Naming me the fifty poems unfinished! (R. Browning)	Example	By stressing there, the word catches the attention of the reader and leads to the emphasis ending with the exclamation point.
'Tis hard to say if greater want of skill Appear in writing or judging ill. (Pope)	Not an Example	Words such as 'tis, to, if, -er, and of are not stressed because they do not carry the message as do the stressed syllables.
Boys in sporadic, tenacious droves Come with sticks, as certainly as Autumn. (Eberhart)	Example	This selection illustrates the masculine line ending wherein the final syllable is stressed.
My mind to me a kingdom is, Such present joys therein I find. (Dyer)	Not an Example	When reading this poem in a natural manner, the second, fourth, sixth and eighth syllables are stressed. To alter this produces a strained, vexing effect that was not intended by the poet.

When you have finished studying the above, please continue to the next page.

Lay a garland on my hearse of the dismal dew, Maidens, willow branches bear, say I died true. My love was false, but I was firm from my hour of birth; Upon my buried body lay lightly, gently earth. (Fletcher)	Example	The second line exemplifies the trochaic meter: MAIDens, WILLow, BRANches, must be read in place of maidENS, wilLOWS, and branCHES.
Nor marble, nor the gilded monuments of princes, shall outlive this powerful rime; But you shall shine more bright in these contents Than unswept stone, besmeared with sluttish time. . (Shakespeare)	Not an Example	Shakespeare generally does not use trochaic meter--this selection is no exception. The beat must be determined by deciding which syllables need the emphasis. This selection does not convey the stronger feeling of trochaic beat.
We stripped and dived and found his body Stuck in the old quicksand. (Seeger)	Example	This line may seem incorrectly marked at first glance--but is trochaic in order to effectively convey the boredom of the soldier telling about the war.
And we just made it out of the Big Muddy With the captain dead and gone.	Not an Example	This example begins with two unstressed syllables and does not incorporate any of the attributes of the trochaic measure.

When you have finished studying the
above, please continue to the next page.

Appendix A-7 Nonrelevant Task Used in Groups 17, 18.

The cry today is for instant improvement of the schools, and there is pressure for wholesale dissemination and development activities without the necessary prior research. The educator from the field invariably asks the university to help him with today's problems, and that is understandable. It is understandable, but more deplorable, that Congress spasmodically lashes out with crisis-oriented legislation but is unsympathetic to balanced, across-the-board, long-range plans. It is tragic that in the U. S. Office of Education the Bureau of Research has thrown its forces heavily on the side of "practical products" and dissemination. While the USOE is a passive patron of basic research, it has done nothing to formulate and sell to Congress a policy that will promote the healthy development of basic investigation.

What the USOE views as research is well illustrated by an early 1966 press release in which it reports enthusiastically on the use of electronically compressed speech to teach the blind, the teaching of first-graders by tape recorders, and teaching third-graders to sing medieval plainsong. Some ad hoc novelties such as these have practical value, and they deserve a trial. But the research program ought to have the higher objective of reexamining educational ideas and the underlying of mental development is true, and what does it imply for educators? How can we account for growth in ability to form elaborate sentence structures? How does motivation for achievement develop? What part does personal identification with the teacher play in forming character and interest? And so on.

Massive dissemination encourages faddism in education. What sounds like a good idea is launched nationwide long before it has been determined that the methods used are really suitable. There is no evidence to justify, for example, the California legislation that requires instruction in foreign language in grades six to eight; the assumptions used to justify the requirements are untested and with the law now a fait accompli, no one is about to test them. The energies of the people who might be giving thoughtful attention to language instruction are diverted into a crash program to write curriculum materials and train teachers The Head Start program is easier to justify, since it reflects a national acceptance of responsibility for the disadvantaged child, and one cannot waste a generation of children while waiting for research. But psychologists do not know whether any of the intervention programs now being installed can produce lasting benefit to intellectual

Continue on Next Page

development; the progrms differ so radically in their assumptions that we can almost be certain that some of them are wrong.

Innovation for innovation's sake is a false value. It crystallizes a practice prematurely and builds up vested interests that discourage hardheaded inquiry and tend to prevent abandonment of the practice when itsglitter wears off.

Now, what recommendations can we make for the proper use of the university as a center for inquiry into education?

First, institutions outside the university should be developed to carry the main burden of demonstration, dissemination, and educational development, The university should, insofar as possible, withdraw from these activities, though it should make the knowledge of its staff available to those who carry them out. Just as aerospace firms and defense laboratories do the developmental work in those engineering fields and the pharmaceutical companies develop products derived from the fundamental studies of the medical school, so educational products will be engineered in institutions resembling Educational Services, Inc., and the Educational Testing Service. These agencies, and the school systems themselves, should do the bulk of in-service training of teachers, though the universities should continue to transmit new ideas to the professional leaders and especially to those who conduct the in-service training.

Second, research should be largely centered in universities, since only the university has the long-range view that permits detached and penetrating inquiry. It is tempting to think of establishing research wings within the development and dissemination institutions, but the hard fact is that in this generation we need to engage every talented researcher as a trainer of researchers and therefore cannot spare him from the university.

Third, the highest priority should be given to recruitment and training of researchers. This calls for breaking down the barriers that now exist between schools of education and other departments. Solid training in one or more of the behavioral, social, and humanistic disciplines is indispensable for thoughtful educational research. Schools of education alone can rarely give that training.

Fourth, we should identify the youngsters who have the greatest

Continue on Next Page

promise as fundamental investigators and should establish the conditions under which they are most likely to become scientific revolutionaries. This means, first of all, financial support for exploratory and unconventional studies, as well as for the neatly canned studies of normal science. It means encouragement of the high-risk activities that do not always pay off, rather than a count-the-publications reward system that locks a man into pedestrian normal science. It means encouraging the man to work on a modest budget that leaves him free to think about his own data.

The improvement of education rests first of all on commitment to the belief that the life of every individual and every nation, and society as a whole, can be lifted to a higher plane of significance through cultivation of the intellect. But improvement will be slight if educational efforts are illuminated by goodheartedness alone. It is a cruel hoax to hail an unsubstantiated method as a cure for an educational deficiency; to adopt it is only to delay the search for underlying causes and for treatments matched to these causes. Intellect begins to play some role in our educational decisions when we test the claims of each new method by assessing its effects in pilot schools. But the intellect takes up its proper duty when it tells us how education and learning proceed, when it tells us why one approach works and another does not, when it identifies the variables that we must adjust to achieve a prescribed effect. The proper mission of the university is to construct, bit by bit, this theory of instruction and of educational systems, while others work on stopgap empirical solutions for educational problems of the moment.

The race between education and catastrophe is not a 60-yard dash, not a matter for spurts of hot-breathed energy. Our generation has a long lap to run. May our pace be strong and our direction sure.

For the person concerned professionally with the improvement of education, this is a time of exhilaration and of despair--exhilaration because our opportunities have expanded to the point where we can almost say that progress is limited only by our capabilities, despair because our capabilities are limiting indeed. Prodigious demands are placed on the school just because it is the one institution under public control that can deliberately cultivate talent and emotional resources--that can on the one hand give individuals the freedom and tools wherewith they carve out a good life, and that can on the other integrate society around principles of opportunity and justice. The nation is ready to back a heroic effort to accomplish these ends, but we professionals do not know enough about learning and instruction to design the desired reform.

Continue on Next Page

While I can only praise in the highest terms the new commitment to education and the enthusiam with which schools are searching for new practices, I am concerned lest the movement may cause the universities, and particularly their schools of education, to neglect their true and unique function. If those whose first calling is the study of education now put off the robe of the scholar and don the armor of the crusader, they will betray the public by leaving the scholar's badly needed work undone.

Effective educational designs, worth careful development and field trial, can emerge only from a deep understanding of learning and motivation.

In the process by which education is improved, we recognize a sequence of activities that starts with basic investigation of the conditions affecting learning, motivation, and instructional effectiveness; carries on through an engineering phase in which practical procedures are designed, tested, and redesigned until they are truly effective, and ends in a marketing phase in which schools are persuaded to adopt the improved methods and teachers are trained to use them. Research, development, dissemination--all three are necessary to keep the educational system moving forward.

Of the three, research is the most difficult to foster. Multiplying research appropriations will not do the job because insightful research requires training and attitudes that are in very short supply. There are few really excellent persons in educational research careers today.

No university can or should fill its faculty entirely with researchers. But any faculty that now contains a modest number of scholars capable of fundamental thought about educational institutions should assign first priority to cultivation of that scholarly capability. With the most careful concentration of existing talent and development of potential talent, we might build up in this country by 1970 a dozen institutions with well-rounded programs of scholarship in education. This will happen, however, only if the better universities hold down their commitments to development and dissemination in order to give first attention to the research mission.

Continue on Next Page

Appendix A-8 Posttest Used In All Groups

The next few pages contain selections of poetry. You are to identify each selection by writing "yes" if you think the selection is and example of trochaic meter poetry, and "no" if you think the selection is not an example of trochaic meter. You may spend as much time per item as is necessary to determine the classification.

When you have finished reading the above, please continue to the next page.

POSTTEST

1. Crouched on the pavement, close by Belgrave Square,
A tramp I saw, ill, moody, and tongue-tied.
(Arnold)

2. Up goes the lark, as if all were jolly
Over the duck-pond the willow shakes.
(Meredith)

3. Break, break, break,
At the foot of thy crags, O Sea!
And I would that my tongue could utter
The thoughts that arise in me.
(Tennyson)

4. Mine eyes have seen the glory of hard work at least.
I have kept the bore unpitted and the action greased.
Even when it ain't a fit night out for man or beast.
(Starbuck)

5. Welling waters winsome word
Wind in warm wan weather.
(Swinburne)

6. Tiger! Tiger! Burning bright
In the forests of the night,
What immortal hand or eye
Could frame thy fearful symmetry?
(Blake)

7. Yet, as if grieving to efface
All vestige of the human race,
On that lone shore loud moans the sea,
But none, alas! shall mourn for me!
(Wilde)

8. When the brokers are roaring like beasts on the floor of the Bourse,
And the poor have the sufferings to which they are fairly accustomed,
And each in the cell of himself is almost convinced of his freedom;
A few thousand will think of this day.
(Auden)

After you have answered all of the above questions, continue to the next page.

9. Since I can never see your face,
And never shake you by the hand,
I send my soul through times and space
To greet you. You will understand.
(Flecker)

10. Oh, to be in England
Now that April's there.
(R. Browning)

11. Fleas
Adam
Had' em
(unknown)

12. Peace is come and wars are over,
Welcome you and welcome all,
While the charger crops the clover
And his bridle hangs in stall.
(Housman)

13. But a young soldier came to our town,
He spoke his mind most candidly.
He asked me quickly to lie down,
And that was very good for me.
(Wickham)

14. Men of England, wherefore plow
For the lords who lay ye low?
Wherefore weave with toil and care
The rich robes your tyrants wear?
(Shelley)

15. Beautiful must be the mountain when ye come,
And bright in the fruitful valleys the streams wherefrom
Ye learn your song:
(Bridges)

16. Somehow--I know not how--as if she ranked
My gift of nine-hundred-years old name
With anybody's gift.
(R. Browning)

After you have finished answering the above questions, continue to the next page.

17. The sun that brief December day
Rose cheerless over hills of gray.
(Whittier)

18. There's a barrel organ caroling across a golden
Street in the city as the sun sinks low;
(Noyes)

19. Therefore he rode and hunted as he might,
Greyhounds he had, swift as a finch in flight;
(Chaucer)

20. Come down, O maid, from yonder mountain height.
What pleasure lives in height (the shepherd sang)
In height and cold, the splendor of the hills?
(Tennyson)

21. He says we are beggars
(Randall)

22. Wild Spirit, which art moving everywhere;
Destroyer and preserver; hear, O, hear!
(Shelley)

23. The God of love my Shepherd is,
And He that doth me feed
While He is mine, and I am His,
What can I want or need?
(Herbert)

24. Spanish waters, Spanish waters, you are ringing in my ears
Like a slow sweet piece of music from the gray forgotten years;
Telling tales and beating tunes, and bringing merry thoughts to me
Of the sandy beach at Muertos, where I would that I could be.
(Masefield)

25. From the ghoulies and the ghosties
Go out ghastly grumblings.
(anonymous)

26. Wee, sleekit, cowin, tim'rous beastie,
Oh, what a panic's in thy breastie!
(Burns)

After you have answered all of the above questions, continue to the next page.

27. Pile the bodies high at Austerlitz and Waterloo.
Shovel them under and let me work--
I am the grass; I cover all.
(Sandburg)

28. Along, alone, alas, he sat.
(unknown)

29. Glory be to God for dappled things--
(Hopkins)

30. My mind to me a kingdom is,
Such present joys therein I find.
(Dyer)

After you have answered all of the above questions, continue to the next page

Appendix A-9 Introduction to Groups 9, 10, 11, 12

The following pages will present to you a program to teach the concept of trochaic meter. Your program will present examples and nonexamples of trochaic meter. Study each program carefully.

When you have finished reading the above, please continue to the next page.

Appendix B Tests
B-1 Answers to Pretest

PRETEST

1. Yes
2. Yes
3. No
4. No
5. Yes
6. No
7. Yes
8. Yes
9. No
10. Yes
11. No
12. No
13. No
14. No
15. Yes
16. Yes
17. No
18. No
19. No
20. No
21. No
22. Yes
23. Yes
24. No
25. Yes
26. Yes
27. Yes
28. No
29. No
30. No

B -2 Answers

POSTTEST

1. No
2. Yes
3. No
4. No
5. Yes
6. Yes
7. No
8. No
9. No
10. Yes
11. No
12. Yes
13. No
14. Yes
15. No
16. No
17. No
18. Yes
19. No
20. No
21. Yes
22. No
23. No
24. Yes
25. Yes
26. No
27. Yes
28. No
29. Yes
30. No

Appendix C-List of Poetry Selections Used In Instance Probability Analysis

1. Never ask of Autumn's falling colors
Where they go; the future days grow duller.
(unknown)

2. Tasks in hours of insight willed
Can be through hours of gloom fulfilled.
(Arnold)

3. Nor marble, nor the gilded monuments
Of princes, shall outlive this powerful rime;
But you shall more bright in these contents
Than unswept stone, besmeared with sluttish time.
(Shakespeare)

4. White were the moorlands and frozen before her,
Green were the moorlands and blooming behind her.
(Kingsley)

5. Pushing through the clouds
Chasing ghosts and shrouds.
(unknown)

6. Let the day perish wherein I was born,
And the night in which it was said.
There is a man child conceived,
Let that day be darkness
(Job: Bible)

7. Says he, "Dear James, to murder me
  Were a foolish thing to do,
For don't you see that you can't cook me,
  While I can--and will--cook you!"
(Gilbert)

8. My mind to me a kingdom is
Such present joys therein I find.
(Dyer)

9. And we just made it out of the Big Muddy
With the captain dead and gone.
(Seeger)

10. Soft and easy is thy cradle
Course and hard thy Savior lay.
(unknown)

11. I will go up to the mountain
And there I will light a fire.
(Austin)

12. Through the noises of the night
She floated down to Camelot:
(Tennyson)

13. The smiles that win, the tints that glow,
But tell of days in goodness spent,
(Byron)

14. Maid of Athens, were we part,
Give, oh give me back my heart!
(Byron)

15. Out of fiendship came the Redman
Teaching settlers where the deer ran.
(Imitation, Hiawatha)

16. He says we are beggars.
(Randall)

17. Glory be to God for dappled things--
(Hopkins)

18. Have I not passed thee on the wooden bridge,
Wrapt in thy cloak and battling with the snow,
They fact toward Hinksey and its wintery ridge?
(unknown)

19. All the hapless silent lovers,
All the prisoners in the prisons,
All the righteous and the wicked,
All the joyous, all the sorrowing,
All the living, All the dying,
Pioneers! O pioneers!
(Whitman)

20. Ruth and Naomi gathered corn and wheat stalks.
They are not afraid but feel secure.
God and Boaz watch their walks,
Taking care that they find plenty near.
(unknown)

21. Motherly, Fatherly, Sisterly, Brotherly!
(unknown)

22. But the wind can so easily whip
The still water to foam,
And the harmless bay waves turn
To storming gray boulders that pound.
(unknown)

23. Take her up tenderly,
Lift her with care;
(Hood)

24. How hast thou merited--
Of all man's clotted clay the dingiest clot?
Alack, thou knowest not
How little worthy of any love thou art!
(Thompson)

25. If the heart of man is depressed with cares,
The mist is dispell'd when a woman appears.
(Gay)

26. Out of childhood into manhood
Now had grown my Hiawatha.
(Longfellow)

27. The God of love my Shepherd is,
And He that doth me feed,
While He is mine, and I am His,
What can I want or need?
(Herbert)

28. When I was one-and-twenty
I heard a wise man say,
"Give crowns and pounds and guineas,
But not your heart away."
(Housman)

29. Give every man thy ear, but few thy voice.
This above all: To thine own self be true.
(Shakespeare)

30. "Spanish ships of war at sea! We have
sighted fifty-three!"
(Tennyson)

31. Heartless, Hopeless, without feeling.
(unknown)

32. Pansies, lilies, kingcups, daisies.
(Wordsworth)

33. He says we are beggars and I say
When it comes to love we are orphans
We are all misplaced or displaced
Persons from another war.
(Randall)

34. Come to the crag where the beacon is blazing,
Come with the buckler, the lance, and the bow.
(Scott)

35. Where are the songsof Spring? Ay, where are they?
Think not of them, thou hast thy music too, --
(unknown)

36. Now the day is over,
  Night is drawing nigh
Shadows of the evening
  Steel across the sky.
(Sabine Bariag-Gould)

37. We stripped and dived and found his body
Stuck in the old quicksand.
(Seeger)

38. Since I can never see your face,
And never shake you by the hand,
I send my soul through times and space
To greet you. You will understand.
(Flecker)

39. Oh, to be in England
Now that April's there.
(R. Browning)

40. My father, he was a mountaineer,
His fist was a knotty hammer;
He was quick on his feet as a running deer,
And he spoke with a Yankee stammer.
(Benet)

41. Peace is come and wars are over,
Welcome you and welcome all,
While the charger crops the clover
And his bridle hangs in stall.
(Housman)

42. But a young soldier came to our town,
He spoke his mind most candidly.
He asked me quickly to lie down,
And that was very good for me.
(Wickham)

43. Men of England, wherefore plow
For the lords who lay ye low?
Wherefore weave with toil and care
The rich robes your tyrants wear?
(Shelley)

44. Beautiful must be the mountain whence ye come,
And bright in the fruitful valleys the streams where from
Ye learn your song;
(Bridges)

45. Somehow--I know not how--as if she ranked
My gift of nine-hundred-years-old name
With anybody's gift.
(R. Browning)

46. The sun that brief December day
Rose cheerless over hills of gray.
(Whittier)

47. There's a barrel organ caroling across a golden
Street in the city as the sun sinks low;
(Noyes)

48. Crouched on the pavement, close by Belgrave Square
A tramp I saw, ill, moody, and tongue-tied.
(Arnold)

49. Up goes the lark, as if all were folly
Over the duck-pond the willow shakes.
(Meredith)

50. Break, break, break
At the foot of thy crags, O Sea!
And I would that my tongue could utter
The thoughts that arise in me.
(Tennyson)

51. Mine eyes have seen the glory of hard work at least.
I have kept the bore unpitted and the action greased.
Even when it ain't a fit night out for man or beast.
(Starbuck)

52. Welling waters winsome word
Wind in warm wan weather.
(Swinburne)

53. Tiger! Tiger! Burning bright
In the forests of the night,
What immortal hand or eye
Could frame thy fearful symmetry?
(Blake)

54. Yet, as if grieving to efface
All vestige of the human race,
On that lens where loud means the sea,
But none, alas! shall mourn for me!
(Wilde)

55. When the breakers are roaring like beasts on the floor
of the Bourse,
And the poor have the sufferings to which they are
fairly accustomed,
And each in the cell of himself is almost convinced
of his freedom;
A few thousand will think of this day.
(Auden)

56. Therefore he rode and hunted as he might,
Greyhounds he had, swift as a finch in flight;
(Chaucer)

57. Come down, O maid, from yonder mountain height.
What pleasure lives in height (the shepherd sang)
In height and cold, the splendor of the hills?
(Tennyson)

58. Spanish waters, Spanish waters, you are ringing in my ears
Like a slow sweet piece of music from the gray forgotten years;
Telling tales, and beating tunes, and bringing weary thoughts to me
Of the sandy beach at Muertee, where I would that I could be.
(Masefield)

59. Wild Spirit, which art moving everywhere;
Destroyer and preserver; hear, O, hear!
(Shelley)

60. Maybe I shall find them among the dead.
Hear me, my chiefs,
I am tired. My heart is sad and sick.
(Joseph)

61. Soldier from the wars returning,
Spoiler of the taken town,
Here is case that asks not earning;
Turn you in and sit you down.
(Housman)

62. From ghoulies and ghosties
Go out ghastly grumblings.
(anonymous)

63. Wee, sleekit, cowin, tim'rous beastie,
Oh, what a panic's in thy breastie!
(Burns)

64. Pile the bodies high at Austerlitz and Waterloo.
Shovel them under and let me work--
I am the grass; I cover all.
(Sandburg)

65. Along, Alone, alas, he sat.
(unknown)

66. Lay a garland on my hearse of the dismal dew,
Maidens willow branches bear, say I died true.
My love was false, but I was firm from my hour of birth;
Upon my buried body lay lightly, gently, earth.
(Fletcher)

67. When they shot Malcolm Little down
On the stage of the Audubon Ballroom,
When his life ran out through bullet holes
(Like people running out when the murder began)
His blood soaked the floor.
(Patterson)

68. I am monarch of all I survey.
(Cowper)

69. Thou who wilt not love do this;
Learn of me what Woman is.
(Herrick)

70. Tis hard to say if greater want of skill
Appear in writing or in judging ill.
(Pope)

71. Could I but live again
Twice my life over,
Would I meet strive again?
(R. Browning)

72. Wailing, Wailing, Wailing, the wind over
land and sea--
(Tennyson)

73. Fleas,
Adam
Had' em
(anonymous)

74. Sure solacer of human cares,
And sweeter hope, when hope despairs!
(Bronte)

75. I want to know
What is really
Going on.
(Resendez')

76. A crack was in the glass
(unknown)

77. The God of love my Shepherd is,
And He that doth me feed,
While He is mine, and I am His,
What can I want or need?
(Herbert)

78. There they are, my fifty men and women,
Naming me the fifty poems unfinished!
(R. Browning)

79. Memory tells me of the many times
We were three in a room, trapped
By the all-constricting walls of a
summer shower.
(Lee)

80. Darkness calling; calling.
I fear: You do not.
Pass by, darkness.
Leave me, angels--
Let the daylight return.
(unknown)

81. Boys in sporadic, tenacious droves
Come with sticks, as certainly as Autumn
(Eberhart)

82. The wine of life keeps oozing drop by drop.
(Fitzgerald)

83. The readers of the Boston Evening Transcript
Sway in the wind like a field of ripe corn.
(Eliot)

84. I don't know why she didn't like my saying
that. She gave me her plaintive smile and her
beautiful eyes filled with tears.
(Maugham)

85. Take thought:
I have weathered the storm,
I have beaten out my exile.
(Pound)

86. Once more within the Potter's house alone
I stood, surrounded by the Shapes of Clay.
(Fitzgerald)

87. Come, my Celia, let us prove,
While we can, the sports of love;
Time will not be ours forever.
He, at length, our goods will sever.
(Johnson)

88. Never ask of Autumn's falling colors
Where they go; the future days grow duller.
(unknown)

89. Tasks in hours of insight willed
Can be through hours of gloom fulfilled.
(Arnold)

90. Nor marble, nor the gilded monuments
Of princes, shall outlive this powerful rime;
But you shall shine more bright in these contents
Than unswept stone, besmeared with sluttish time.
(Shakespeare)

91. White were the moorlands and frozen before her.
Green were the moorlands and blooming behind her.
(Kingsley)

92. Pushing through the clouds
Chasing ghosts and shrouds.
(unknown)

C-1 Instance Probability Analysis for Exemplars

Eg

Selection #	Percentage	Selection #	Percentage
14	81%	85	49%
30	82%	81	47%
36	83%	78	47%
41	82%	64	42%
53	80%	49	44%
		17	46%
		7	46%
69	70%		
58	73%		
52	72%	68	36%
43	75%	62	39%
32	74%	60	35%
26	75%	37	37%
19	76%	23	34%
10	77%	20	36%
		16	39%
92	68%		
87	68%	75	25%
80	62%	47	24%
72	60%	35	20%
61	66%	33	22%
39	60%	11	25%
31	69%		
15	66%		
5	64%		
1	68%		
89	50%		
88	58%		
2	59%		

CD104

C-2 Instance Probability Analysis Nonexemplars

Selection #	Percentage	Selection #	Percentage
9	82%	90	50%
22	85%	91	56%
83	82%	76	59%
		74	52%
		73	52%
		70	56%
		66	55%
84	75%	63	52%
79	74%	57	51%
71	74%	56	55%
67	76%	38	58%
55	72%	4	59%
50	70%		
44	76%		
42	73%		
24	70%		
6	76%	86	46%
		65	43%
		46	42%
		28	45%
		27	41%
82	69%	12	47%
59	64%	8	40%
54	66%		
51	66%		
40	64%		
34	62%		
29	69%	77	39%
25	63%	45	31%
21	64%	13	34%
18	65%		
3	65%		

Appendix D-1

Analysis of Variance Table
For Correct Classification

Source	df	MSE	F
Treatments (A)	8	331.38	23.67*
Pretest vs. No Pretest (B)	1	1.09	0.07
A X B	8	12.86	0.92
Error	162	14.00	

*$p < .01$

Appendix D-2

Analysis of Variance Table
For Overgeneralization

Source	df	MSE	F
Treatments (A)	8	248.81	15.57*
Pretest vs. No Pretest (B)	1	23.92	1.49
A X B	8	17.63	1.13
Error	162	15.98	

*$\underline{p} < .01$

Appendix D-3

Analysis of Variance Table
For Undergeneralization

Source	df	MSE	F
Treatments (A)	8	429.57	29.79*
Pretest vs. No Pretest (B)	1	35.82	2.48
A X B	8	28.13	1.95
Error	162	14.42	

*$p < .01$

Appendix D-4

Analysis of Variance Table
For Misconception

Source	df	MSE	F
Treatments (A)	8	64.00	4.83*
Pretest vs. No Pretest (B)	1	13.47	1.01
A X B	8	19.33	1.46
Error	162	13.25	

*$p < .01$

Appendix Study Two

		Page
A	Introduction to all Programs	
B	Introduction to Pretest	
C	Pretest	
D	Answers to Pretest	
E	Introduction to Task	
F	Definition of RX_2 Crystals	
G	Correct Classification Task	
H	Overgeneralization Task	
I	Undergeneralization Task	
J	Misconception Task	
K	Introduction to Nonrelevant Task	
L	Nonrelevant Task	
M	Introduction to Posttest	
N	Posttest	
O	Answers to Posttest	
P	Posttest Probabilities	
Q	Construction of Posttest	
R	Construction of Tasks	
S	Conclusion to Program	
T	Analysis of Variance Tables	

Appendix A Introduction To All Programs

Dear Student:

You have been selected to participate in a research grant funded by the United States Office of Education. The instructions you are to receive have been developed to attain specific information. Please follow all directions carefully.

This program is designed to be self-instructional. The experimenter cannot answer any questions. If for some reason you cannot continue with the program, please take the program to the experimenter and leave quietly. There are not timed breaks so once you begin, continue until finished.

When you have finished reading the above, please continue to the next page.

Appendix B Introduction to Pretest

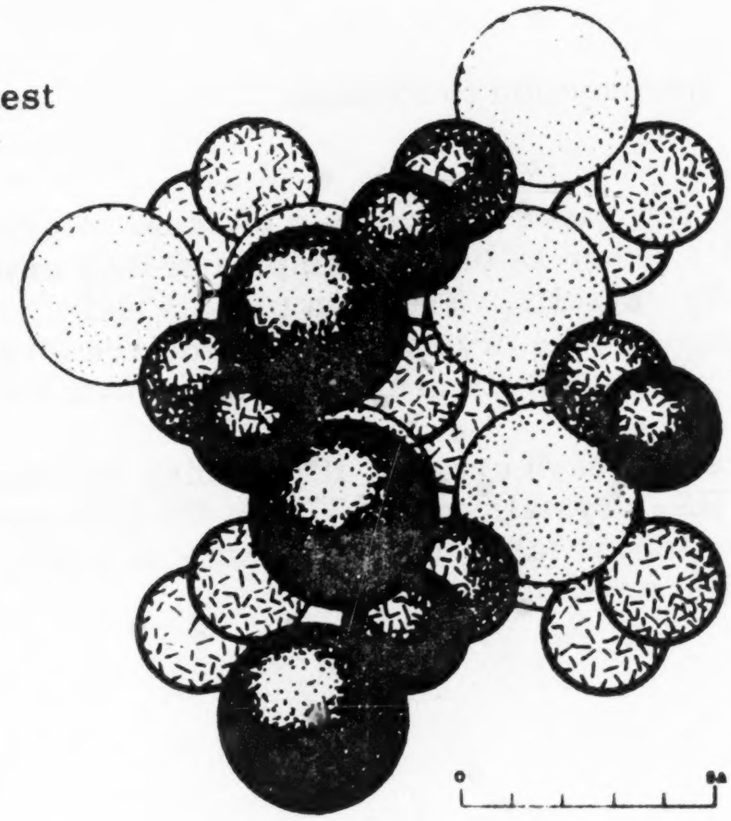
The next few pages will contain pictures of crystals which you are to identify either as being examples of RX_2 crystals or not examples of RX_2 crystals. This is a pretest and it is used to determine your knowledge of RX_2 crystals before you are given instructions. We assume that most students are not familiar with RX_2 crystals.

After viewing each crystal, if you think the item is true (an example of a RX_2 crystal), mark "T" on the IBM answer sheet, if you think the item is false (not an example of a RX_2 crystal), mark "F" on the IBM answer sheet.

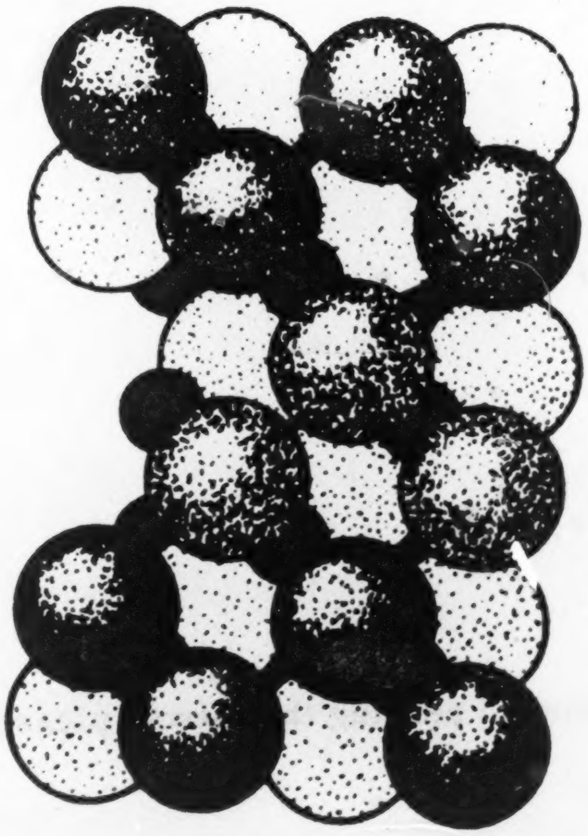
When you have finished reading the above, please continue to the next page.

Appendix C Pretest

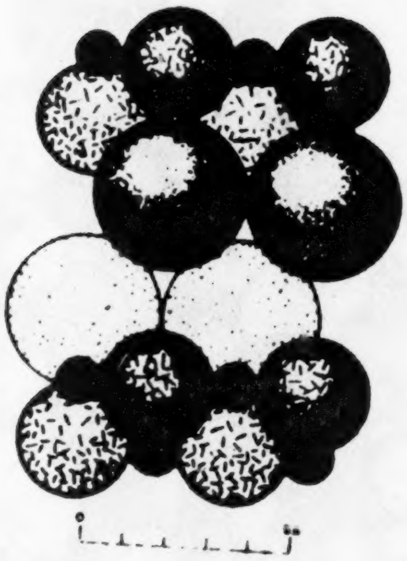
#1



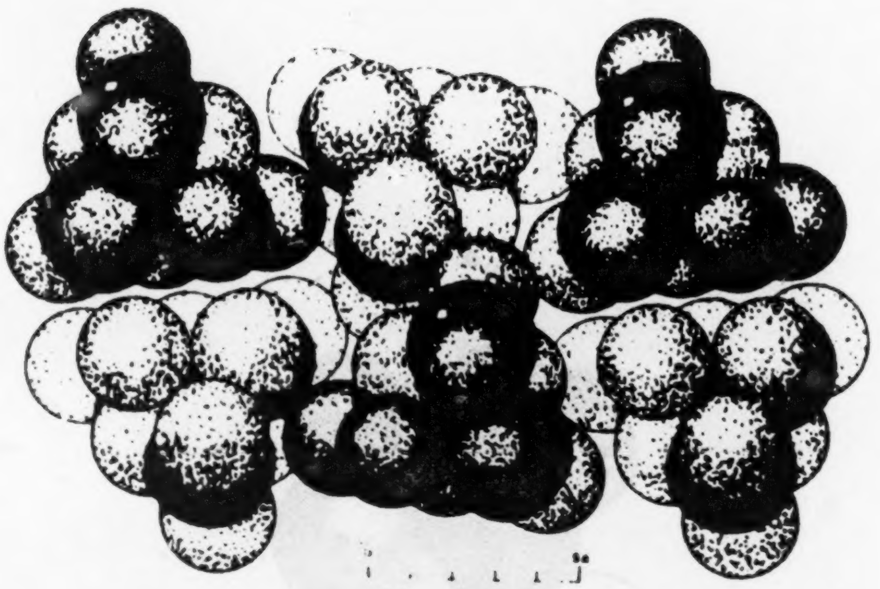
#2



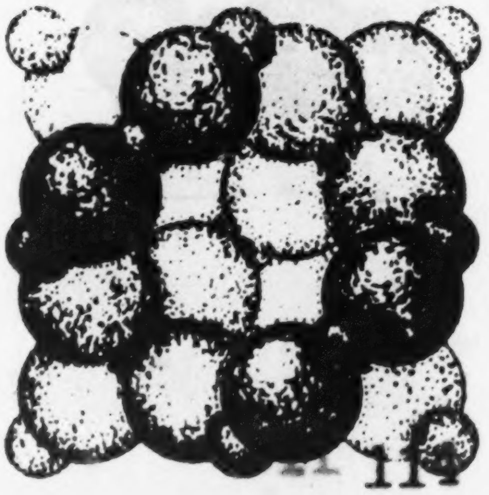
#3



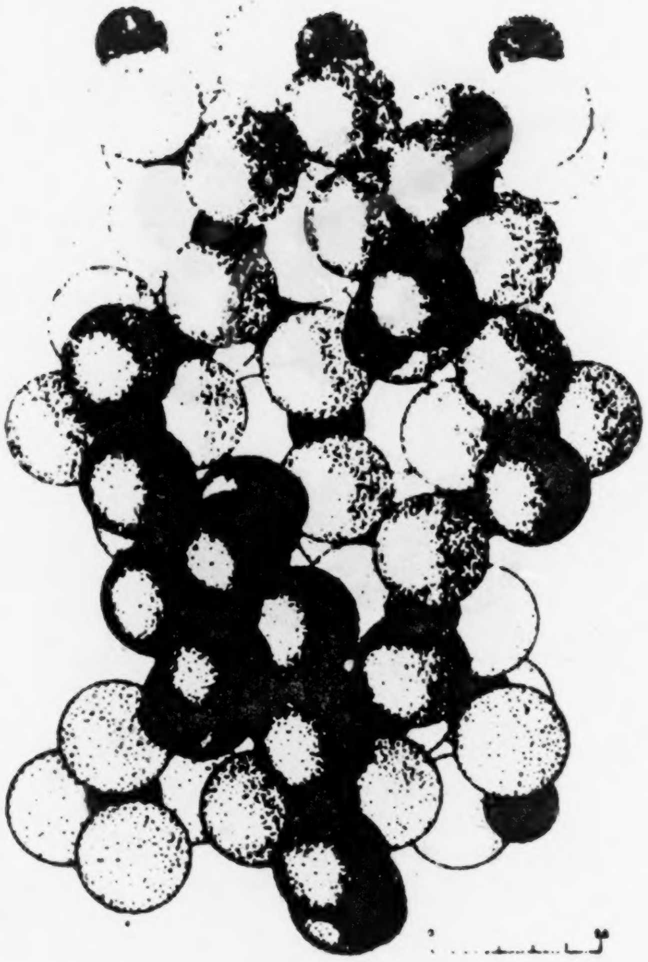
#4



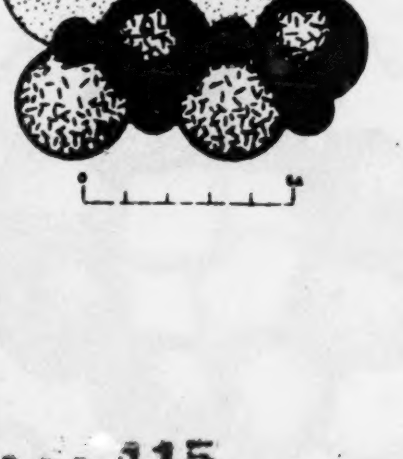
#5



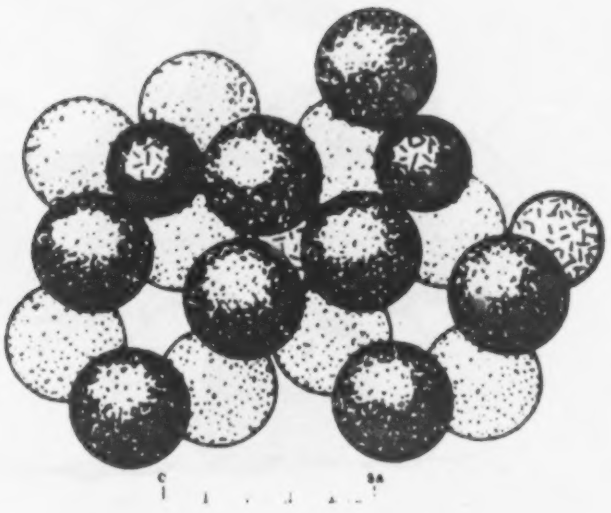
#6



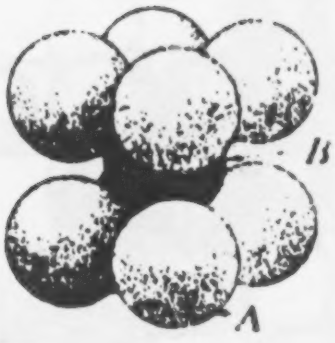
#7



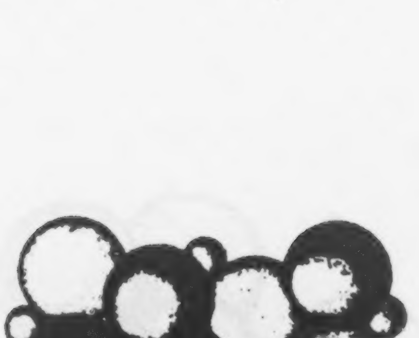
#8



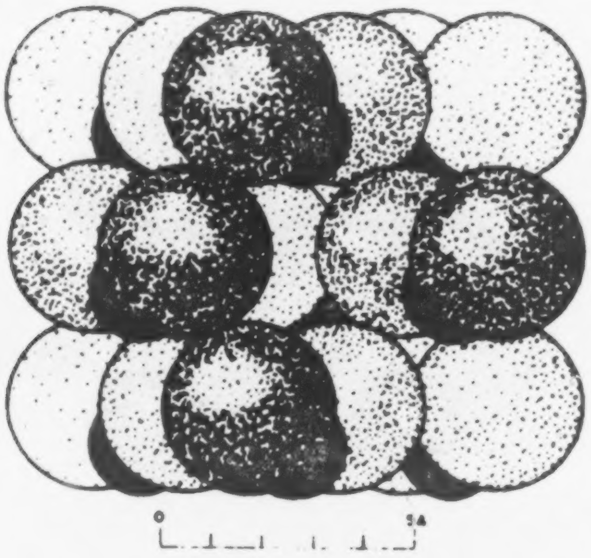
#9



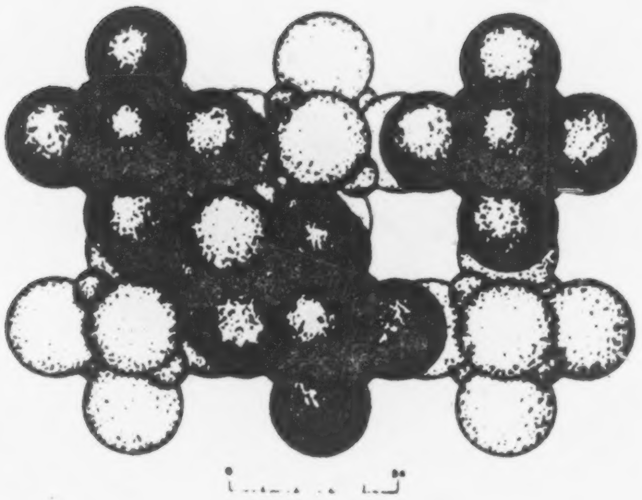
#10



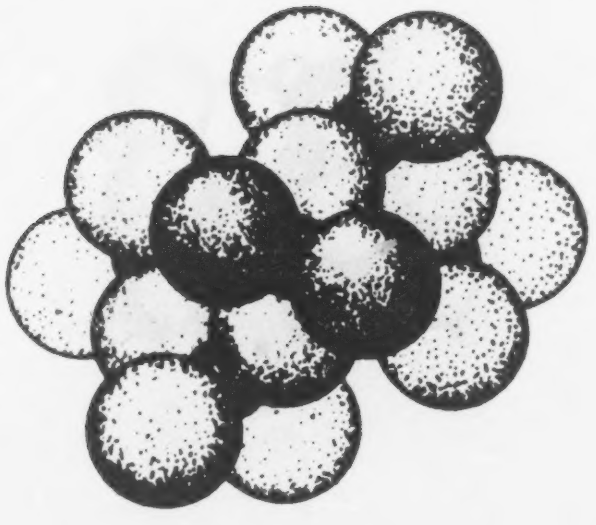
#11



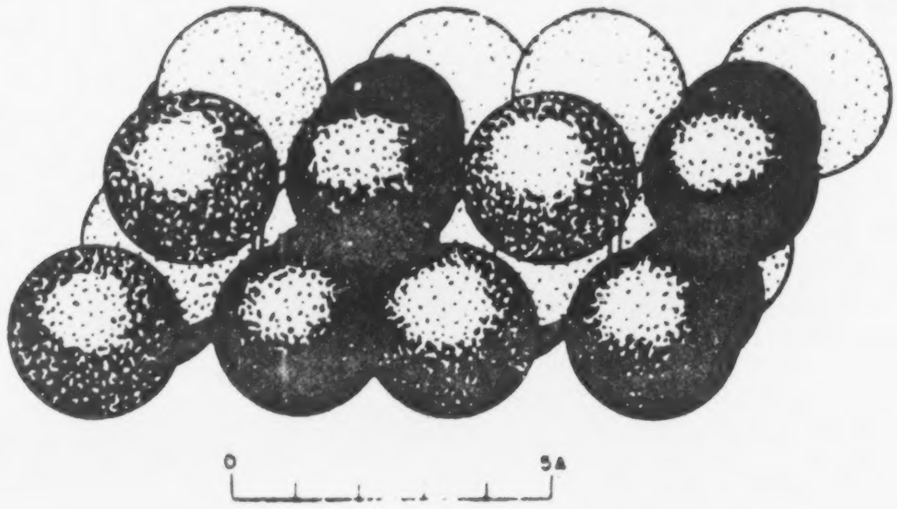
#12



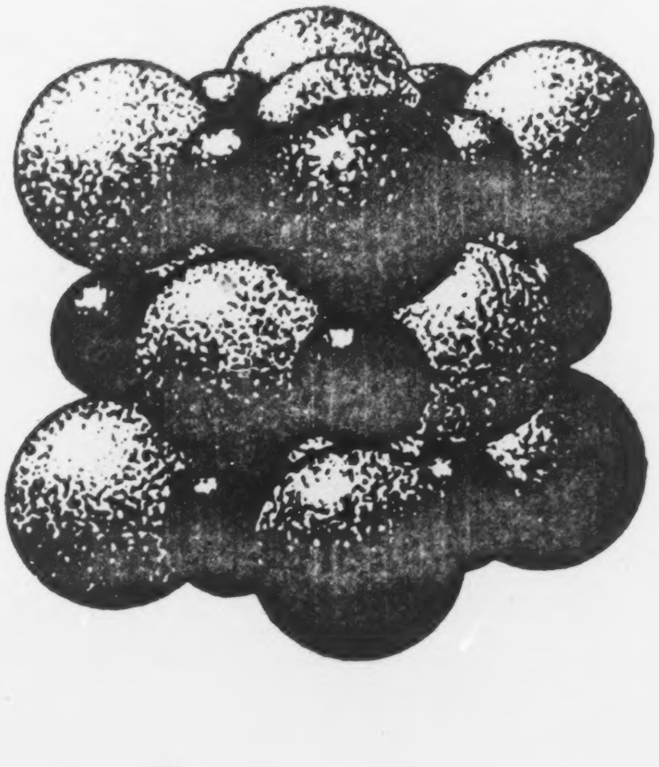
#13



#14



#15



Appendix D Answers to Pretest

1. T
2. F
3. T
4. F
5. T
6. T
7. T
8. F
9. F
10. T
11. T
12. F
13. T
14. T
15. T

Appendix E Introduction to Task

The following pages will present to you a program to teach the concept of RX_2 crystals. The programs vary in length according to a random selection. Study each program carefully.

When you have finished reading the above, please continue to the next page.

Appendix F Definition of RX_2 Crystals

Read the following definition of crystals carefully. You can return to this page during the program. After studying the definition, you will be tested on how well you can identify examples from nonexamples.

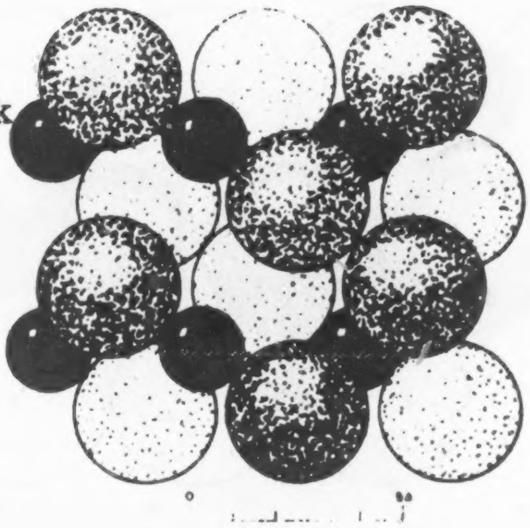
Definition:

Crystals are made up of groups of identical molecules which are comprised of spheres called atoms. The single crystals you are to be tested on may not be complete in and of themselves, but remember that crystals are always symmetrical, so what you don't see may still be present. You must attune yourself to the basic atomic structure or the repeating clusters of atoms. There is a type of crystal called RX_2 which has a two to one ratio in its atomic structure, i.e., for a given atom there will be another two atoms (or clusters of atoms) attached to it in repeating fashion.

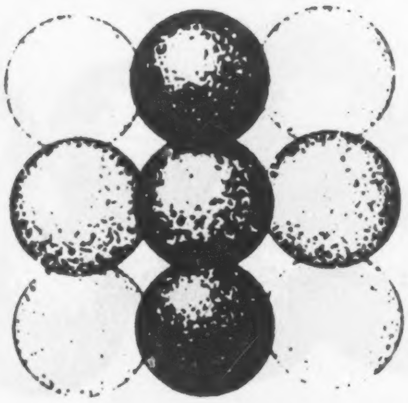
The test item pictures have been shaded to show dimension and depth. You will be shown 8 examples and 8 nonexamples, then you will be tested.

TURN THE PAGE AND CONTINUE WITH THE PROGRAM

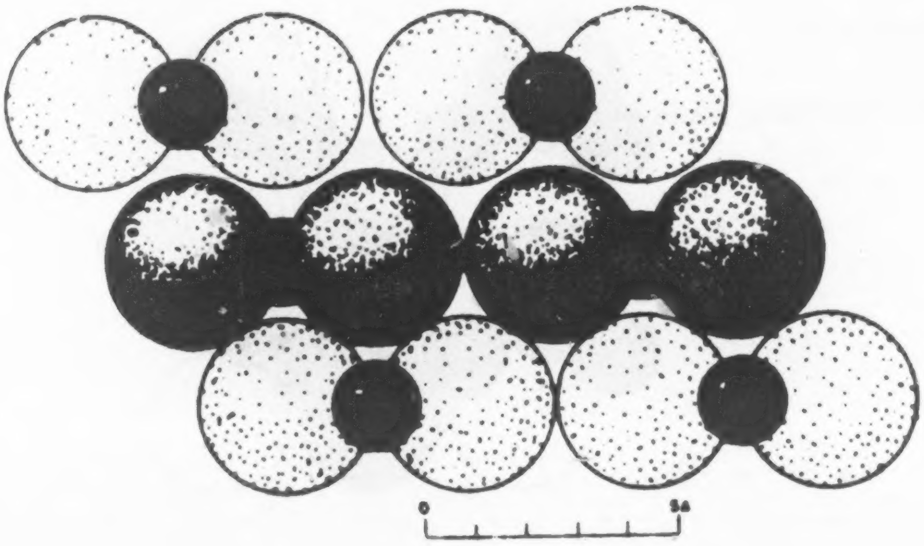
Appendix G
Correct
Classification Task



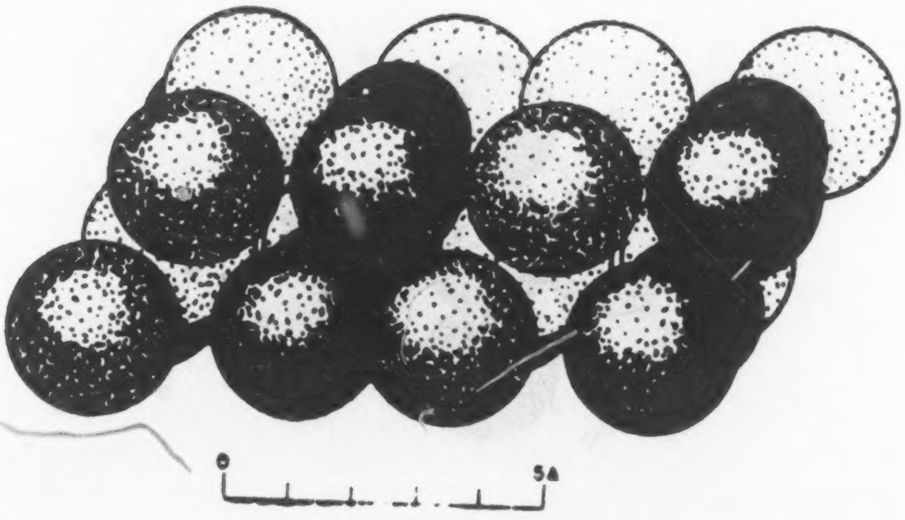
EXAMPLE



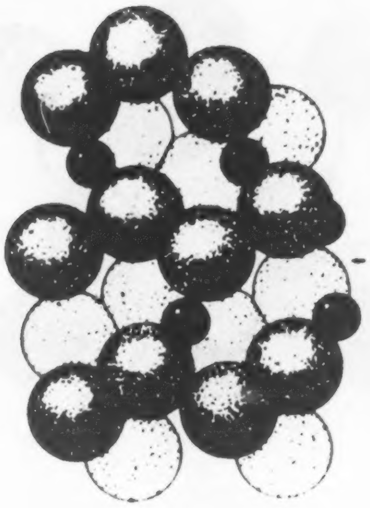
NOT AN EXAMPLE



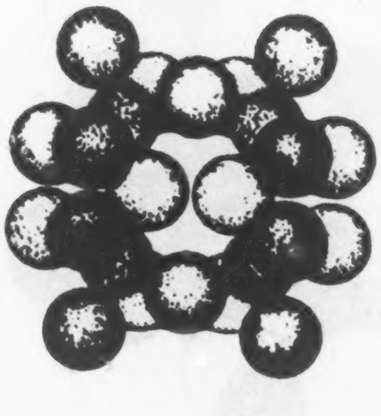
EXAMPLE



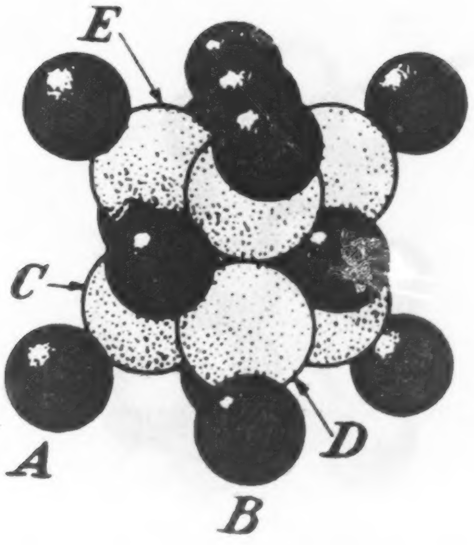
NOT AN EXAMPLE



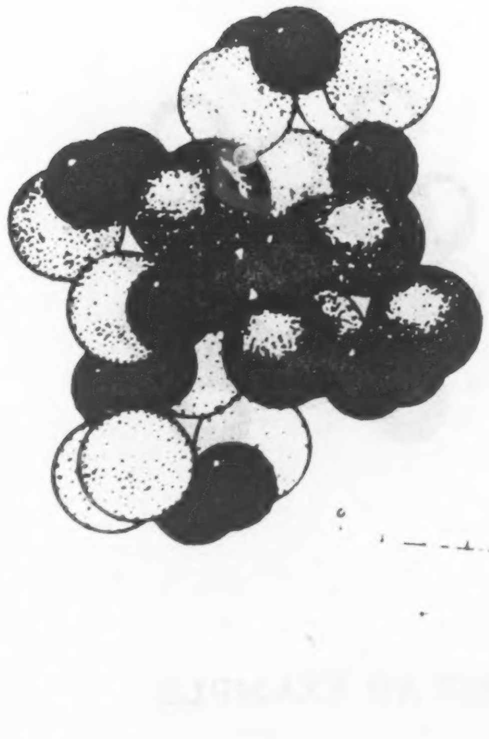
EXAMPLE



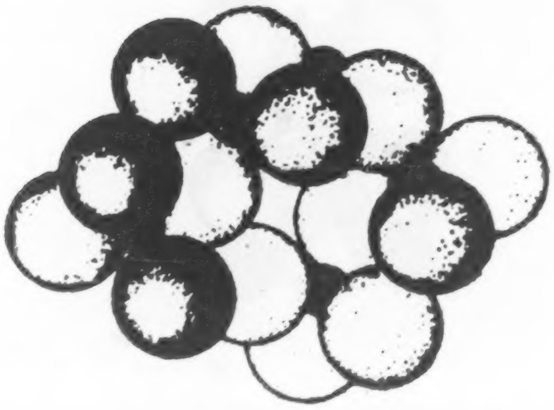
NOT AN EXAMPLE



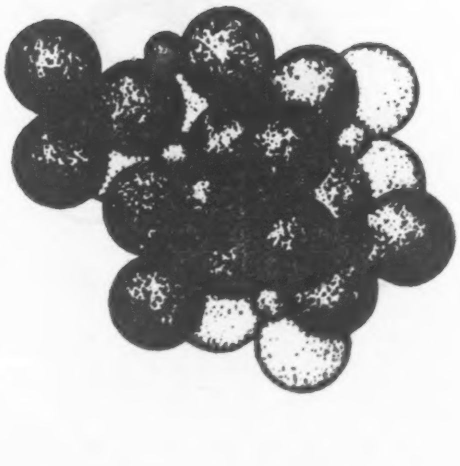
EXAMPLE



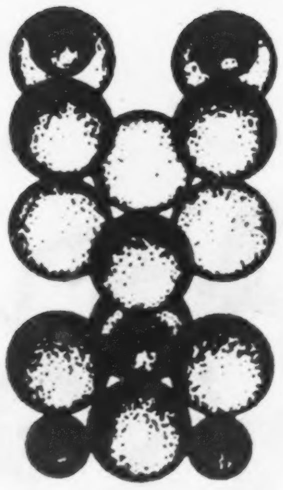
NOT AN EXAMPLE



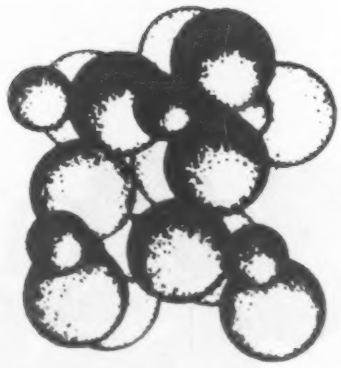
EXAMPLE



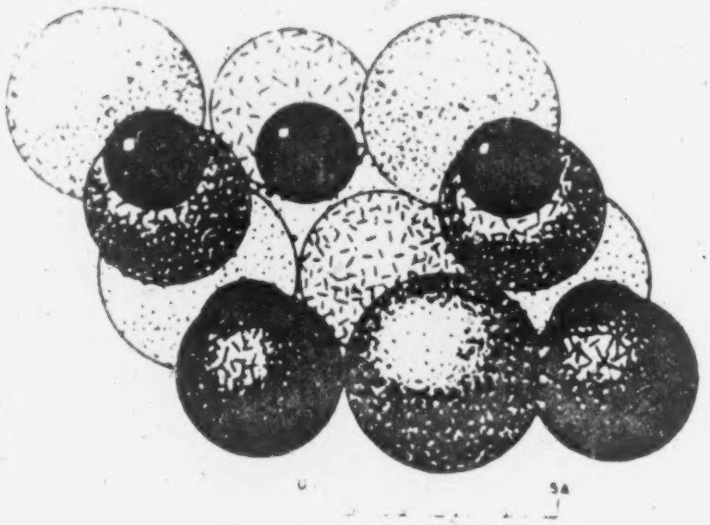
NOT AN EXAMPLE



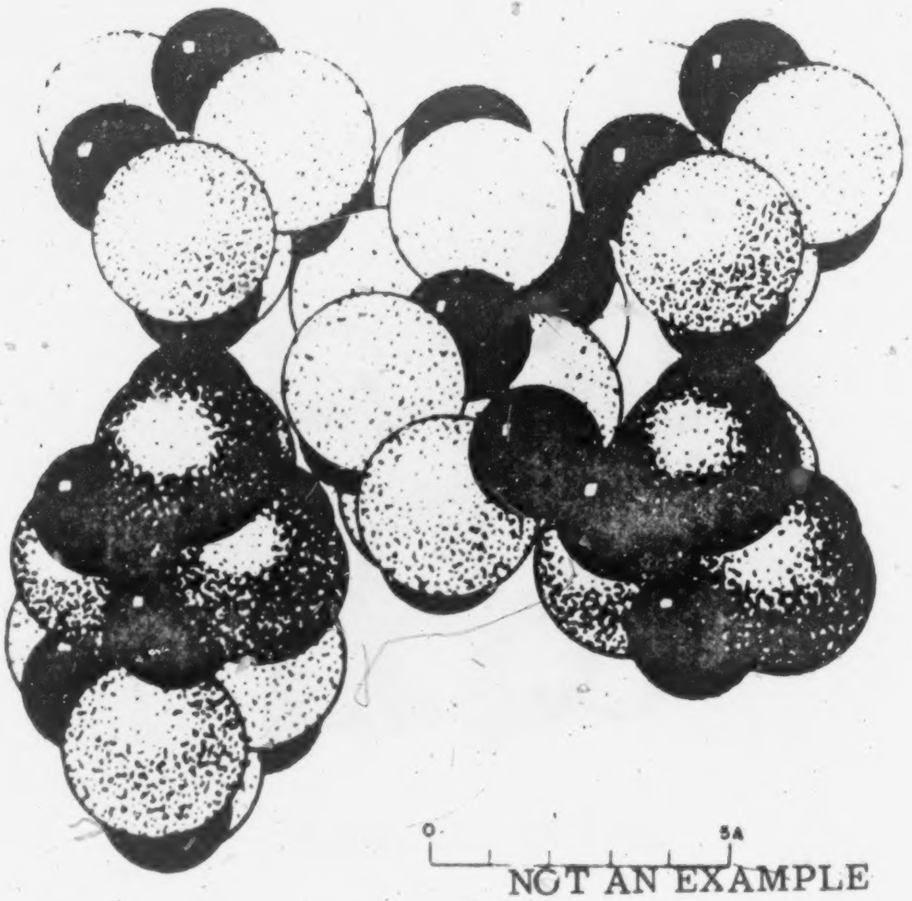
EXAMPLE



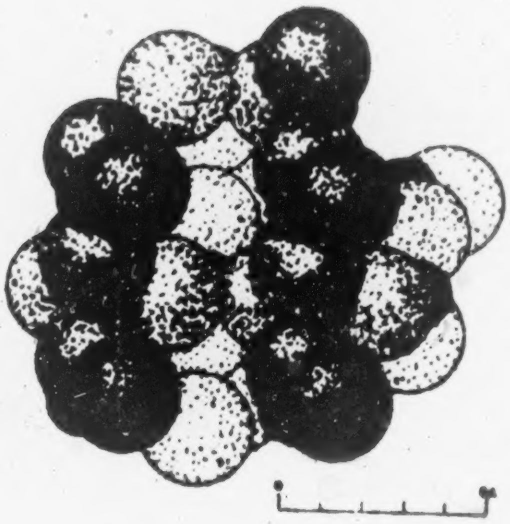
NOT AN EXAMPLE



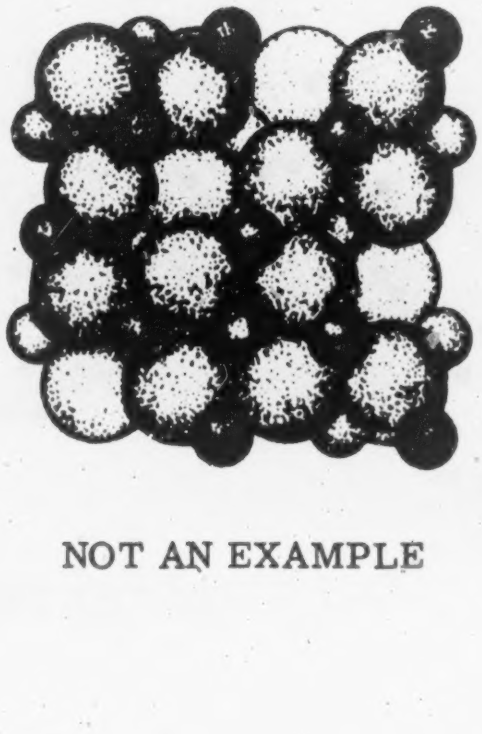
EXAMPLE



NOT AN EXAMPLE

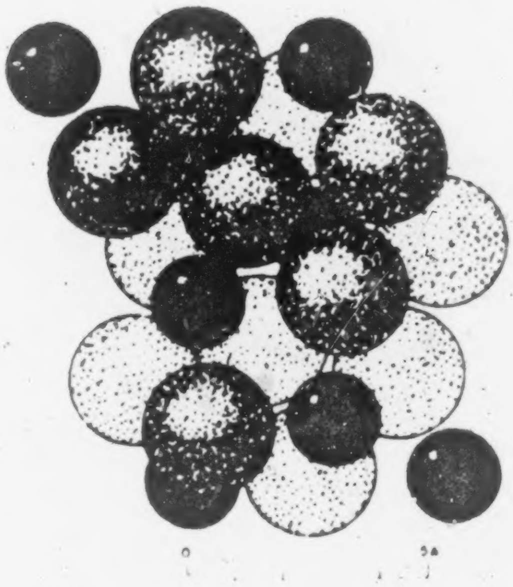


EXAMPLE

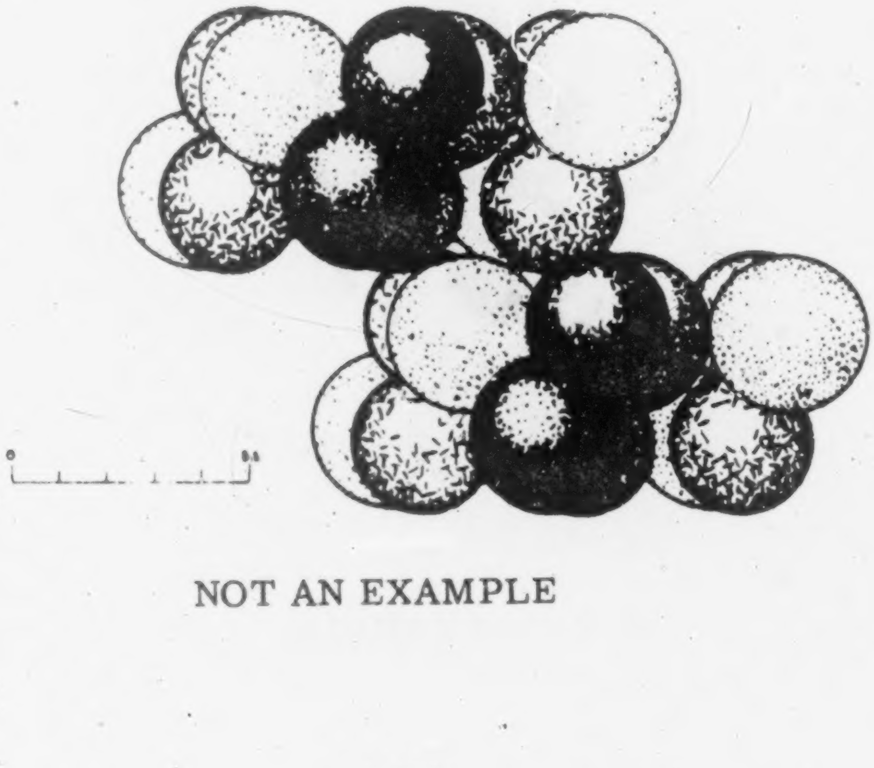


NOT AN EXAMPLE

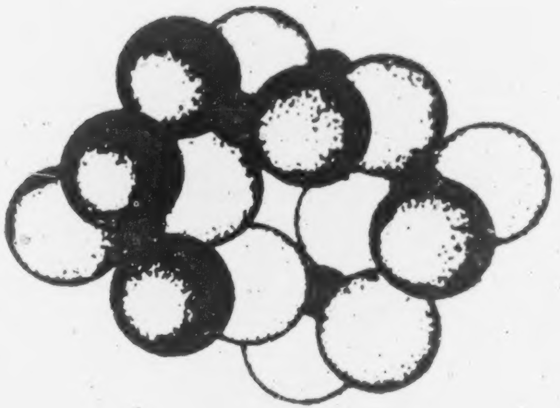
Appendix H
Overgeneralization
Task



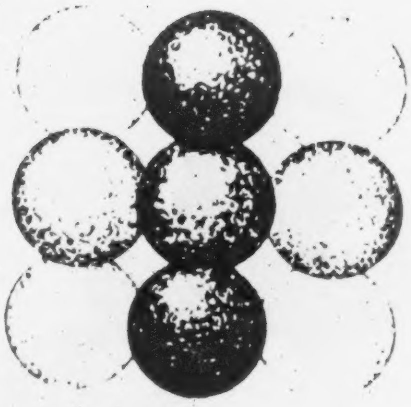
EXAMPLE



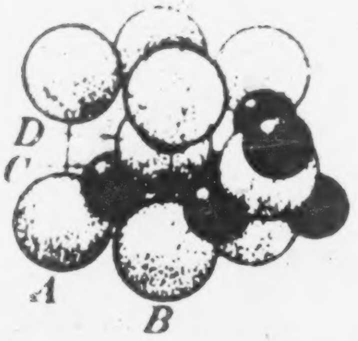
NOT AN EXAMPLE



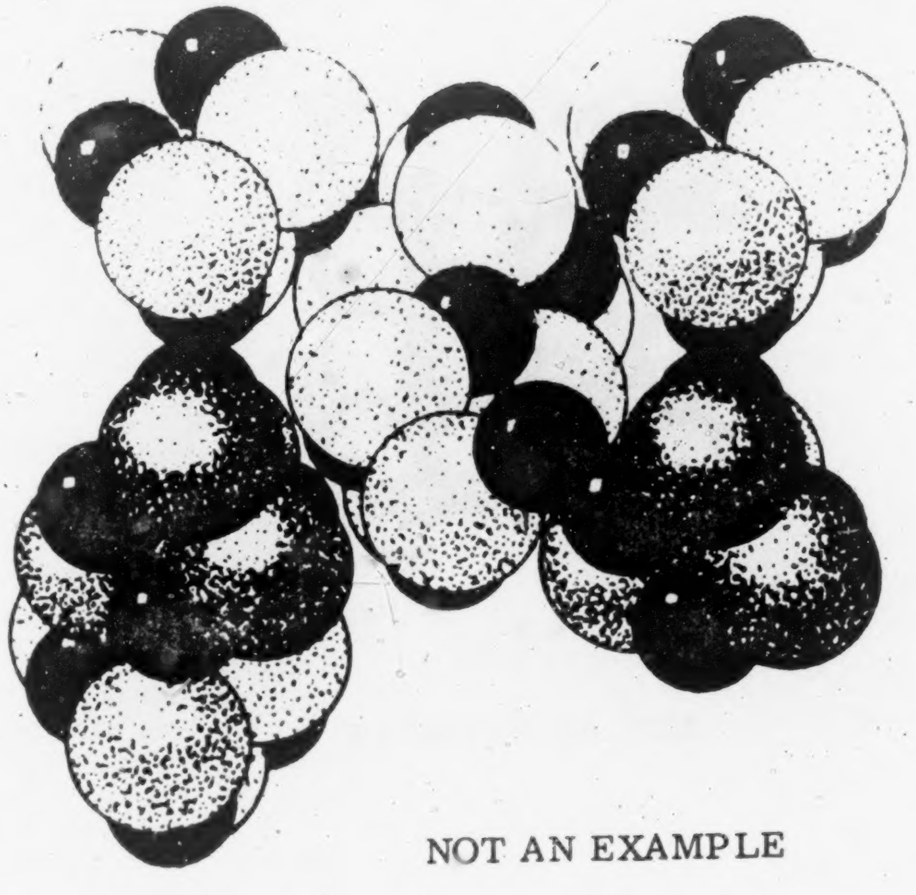
EXAMPLE



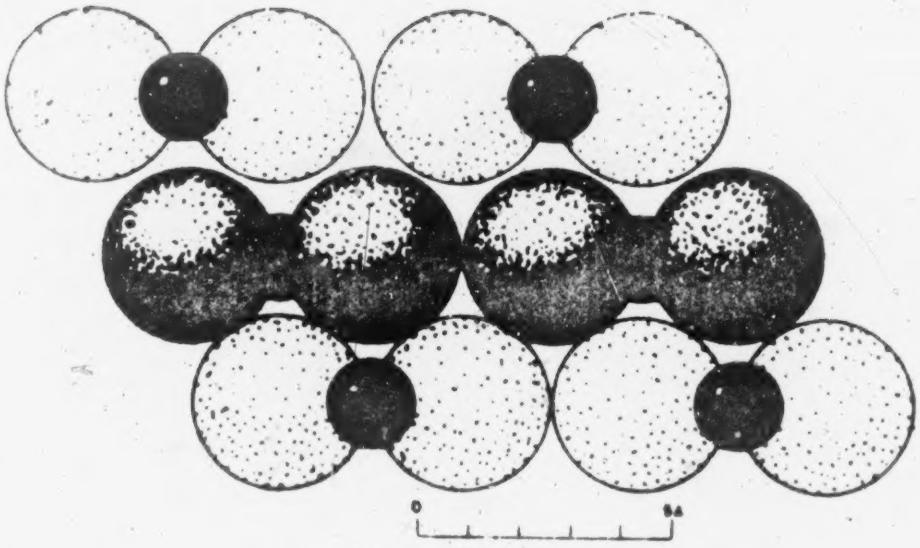
NOT AN EXAMPLE



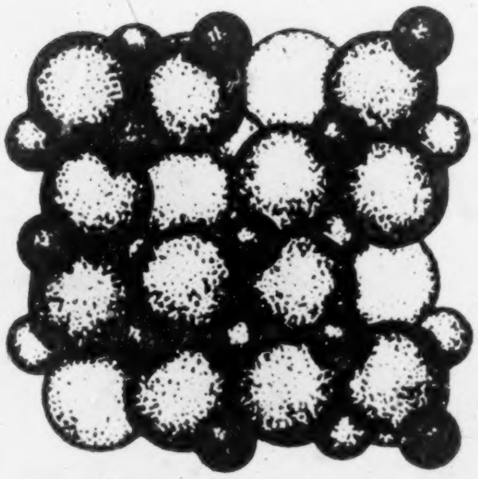
EXAMPLE



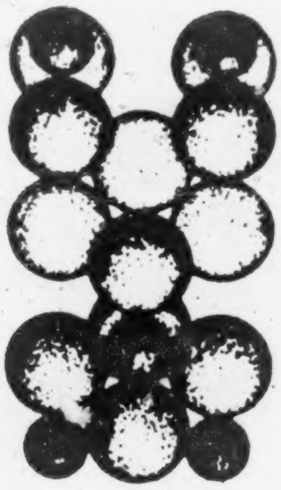
NOT AN EXAMPLE



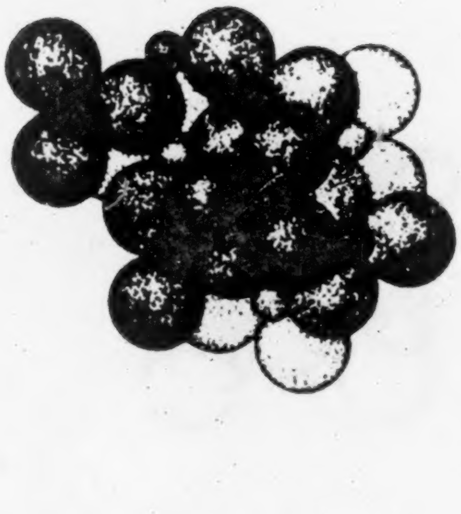
EXAMPLE



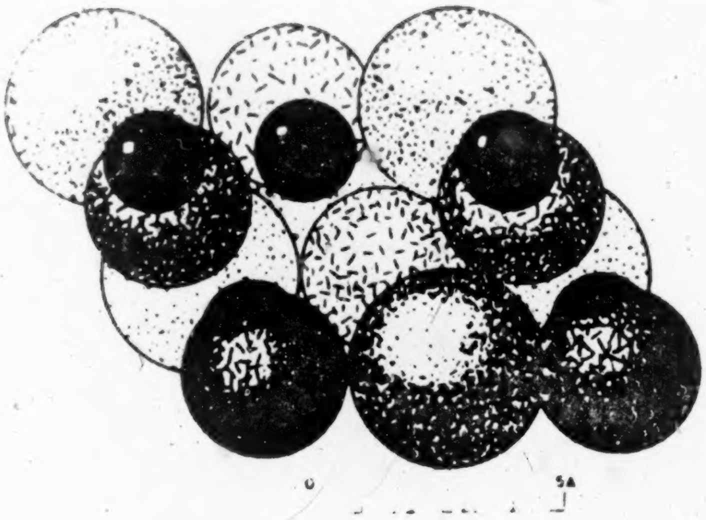
NOT AN EXAMPLE



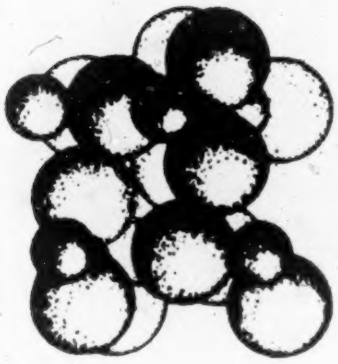
EXAMPLE



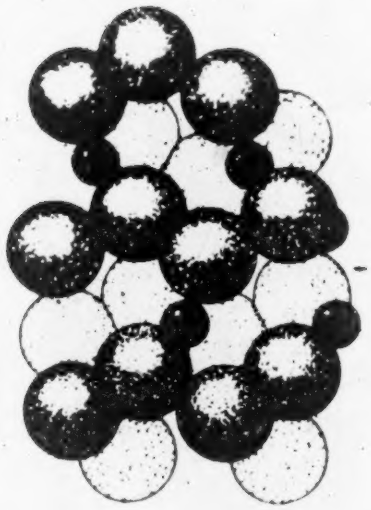
NOT AN EXAMPLE



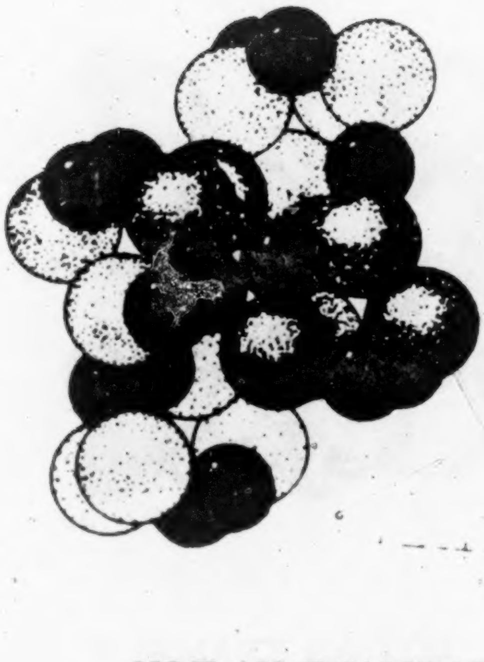
EXAMPLE



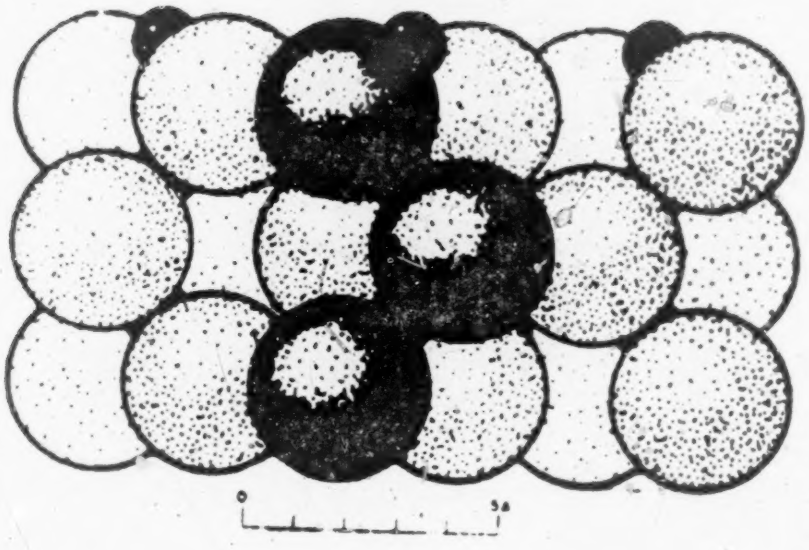
NOT AN EXAMPLE



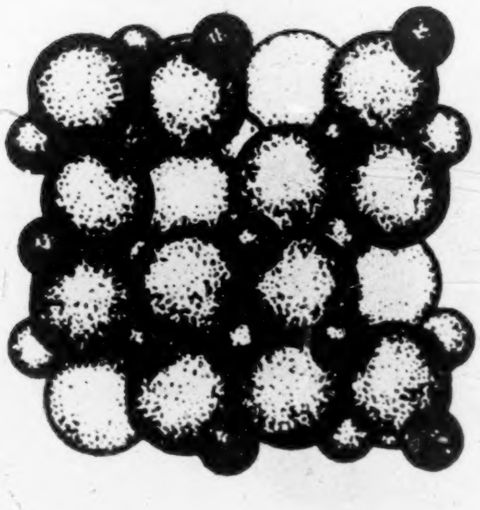
EXAMPLE



NOT AN EXAMPLE

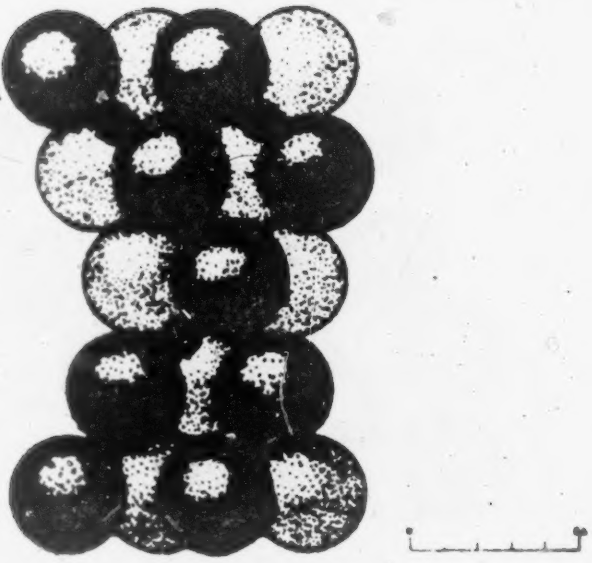


EXAMPLE

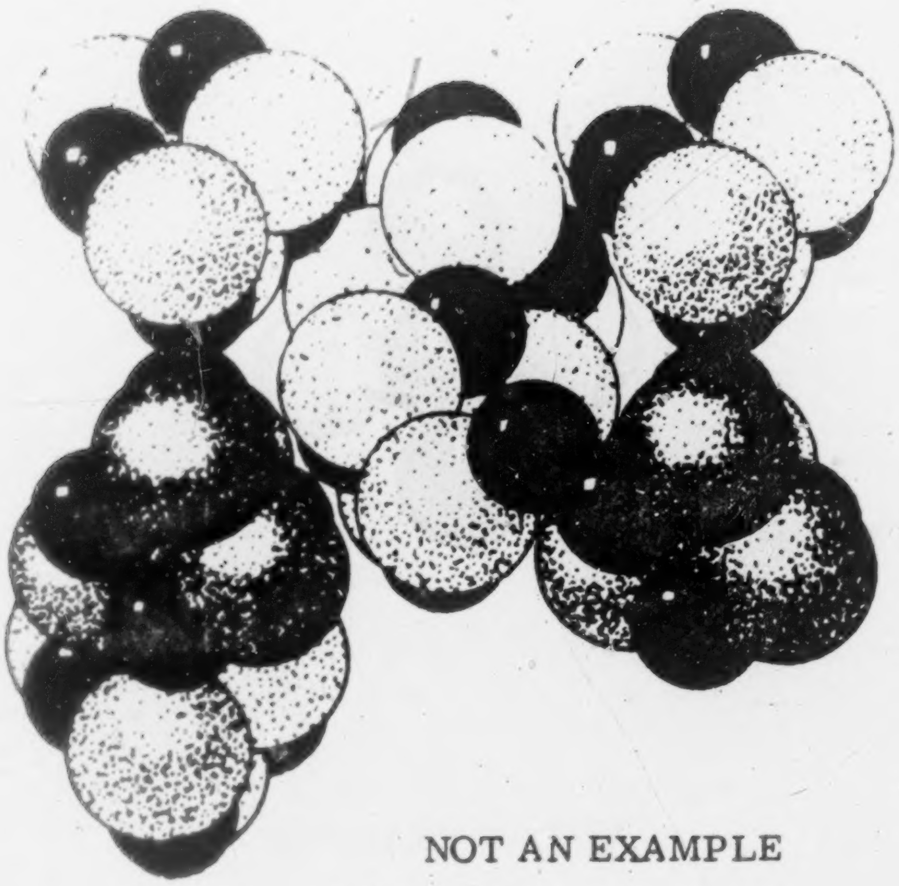


NOT AN EXAMPLE

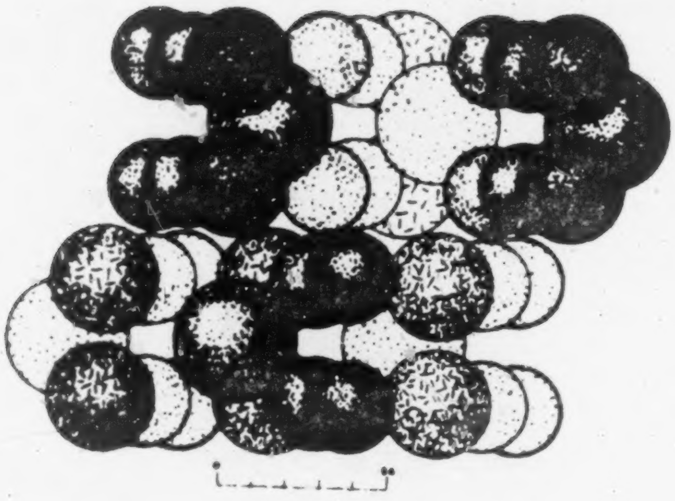
Appendix I
Undergeneralization Task



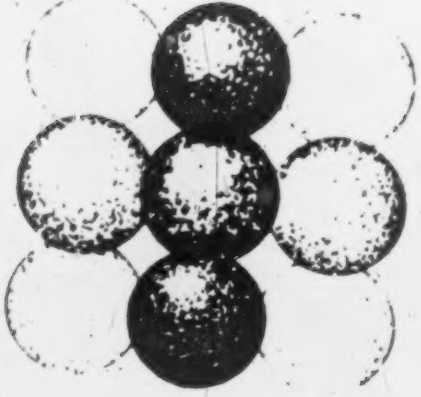
EXAMPLE



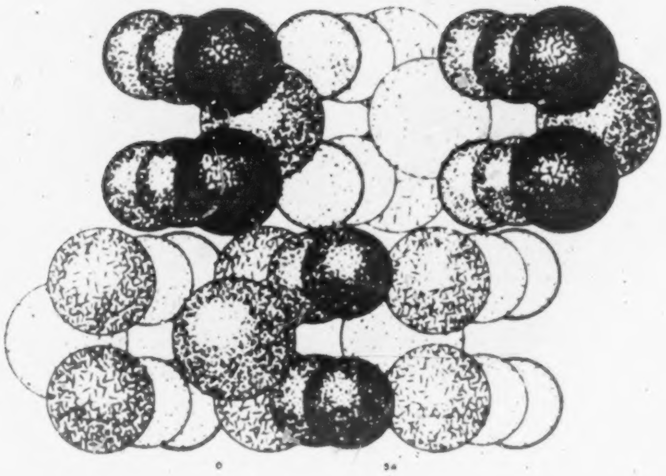
NOT AN EXAMPLE



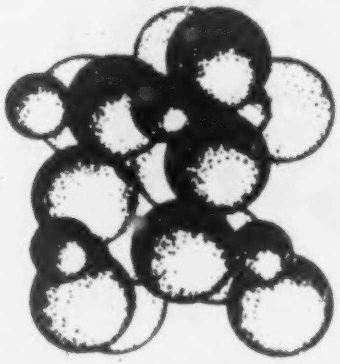
EXAMPLE



NOT AN EXAMPLE

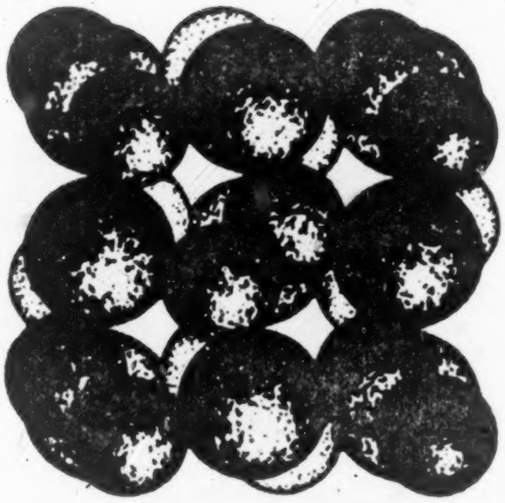


EXAMPLE

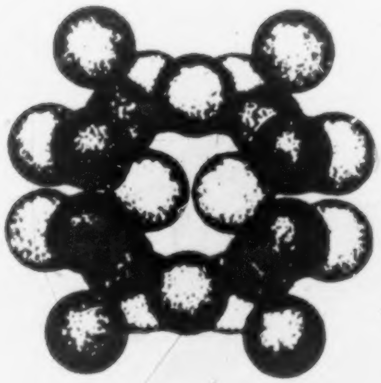


NOT AN EXAMPLE

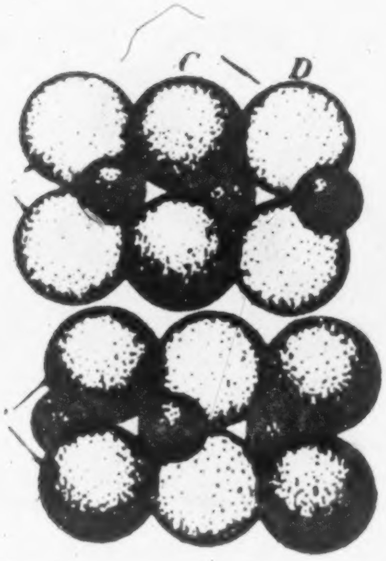
146



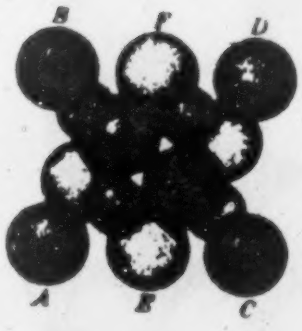
EXAMPLE



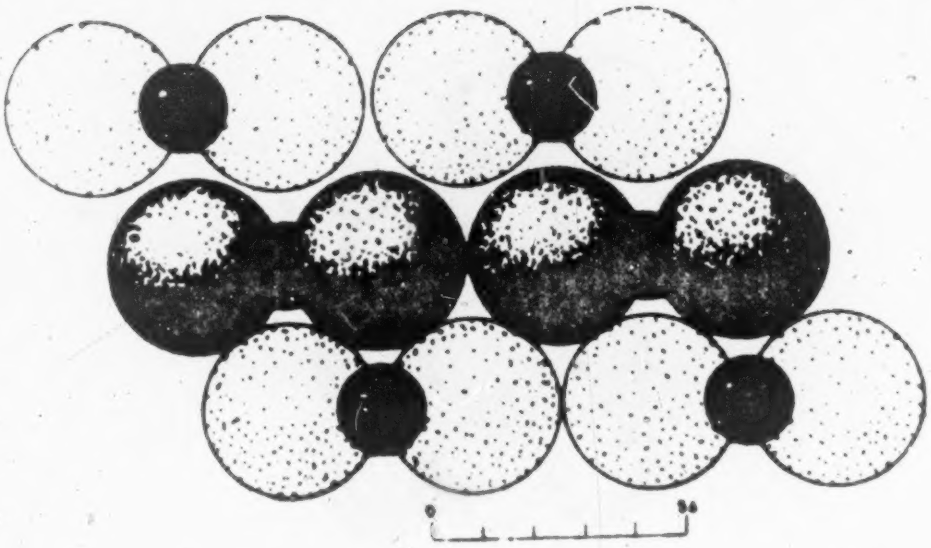
NOT AN EXAMPLE



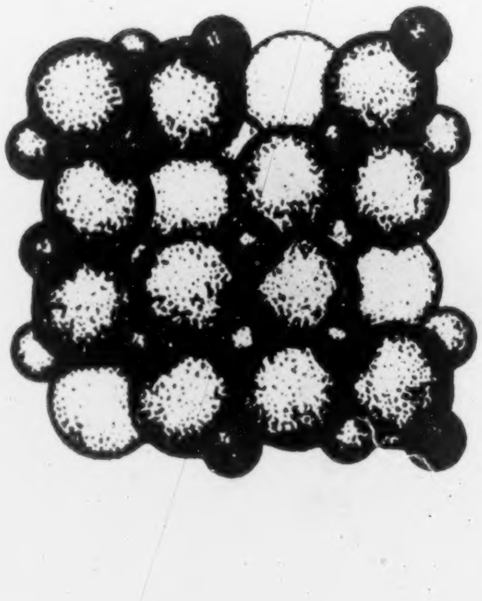
EXAMPLE



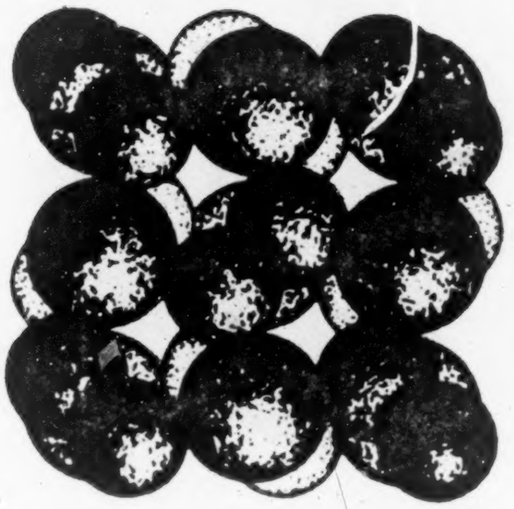
NOT AN EXAMPLE



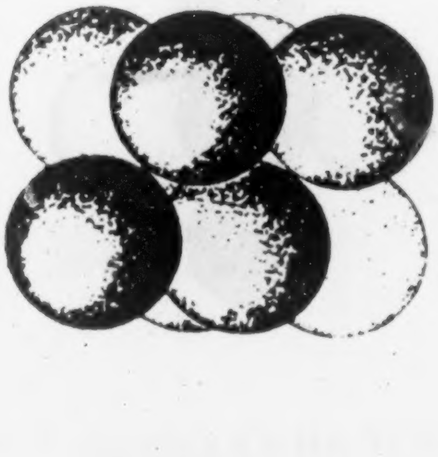
EXAMPLE



NOT AN EXAMPLE

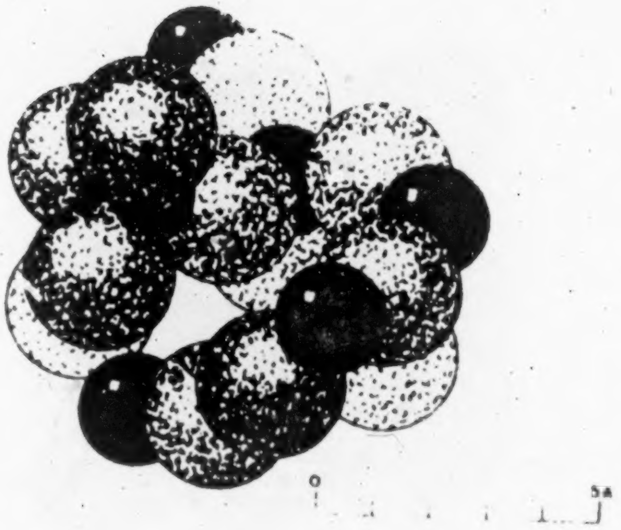


EXAMPLE

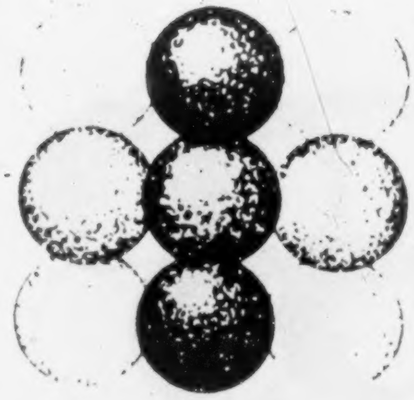


NOT AN EXAMPLE

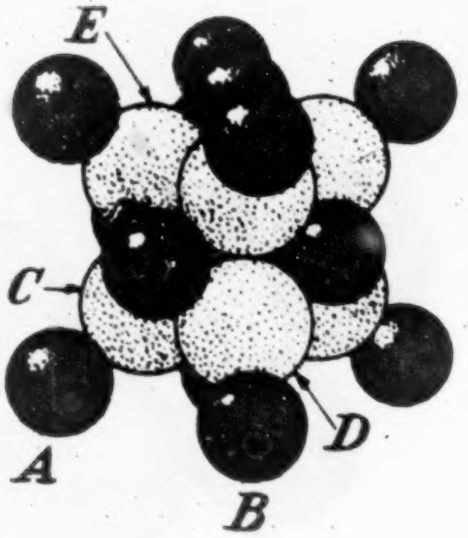
Appendix J
Misconception Task



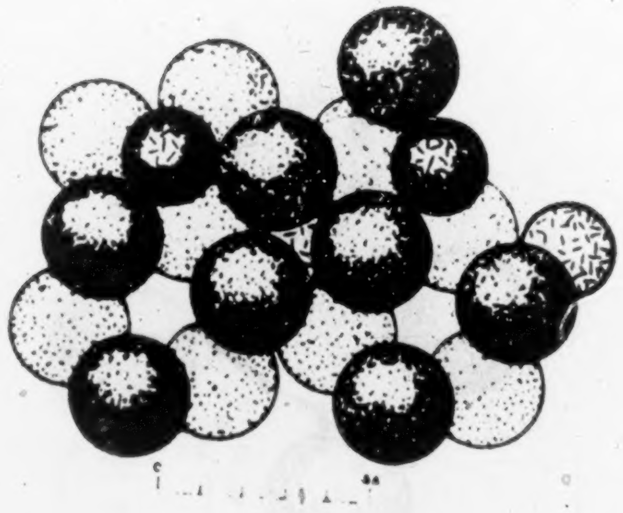
EXAMPLE



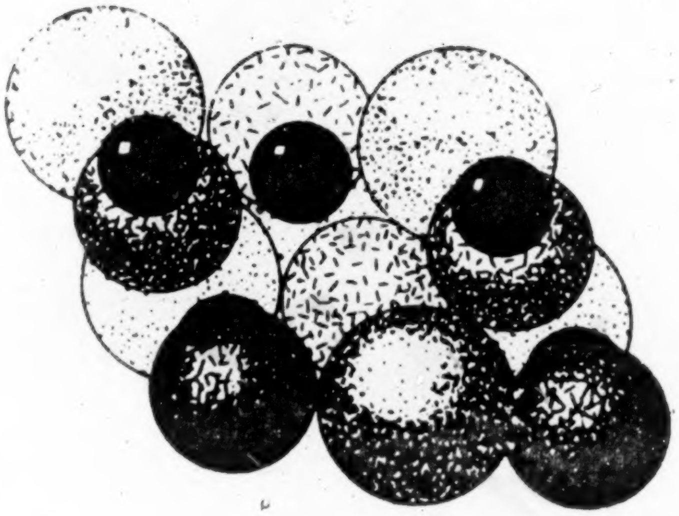
NOT AN EXAMPLE



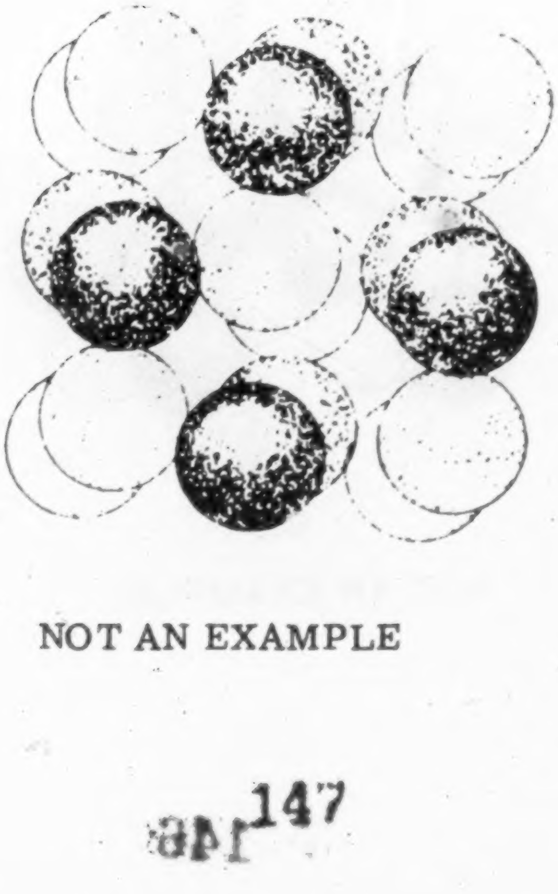
EXAMPLE



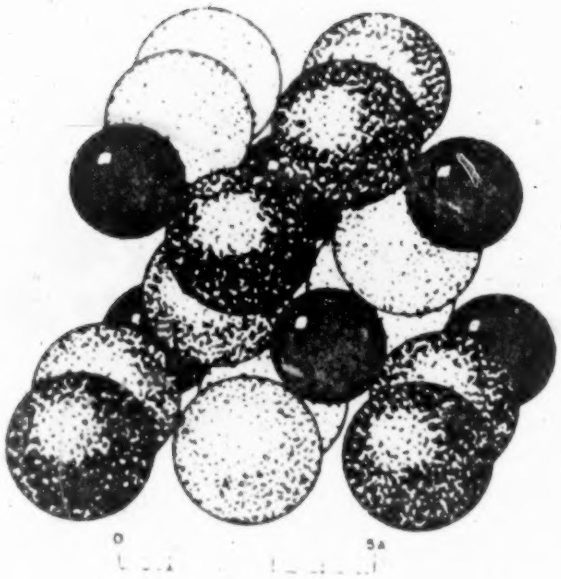
NOT AN EXAMPLE



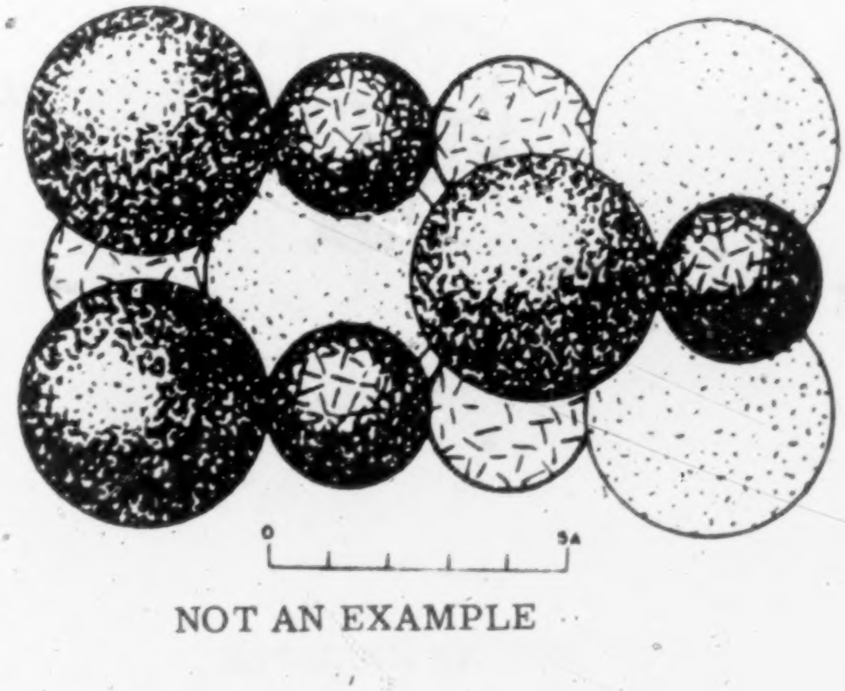
EXAMPLE



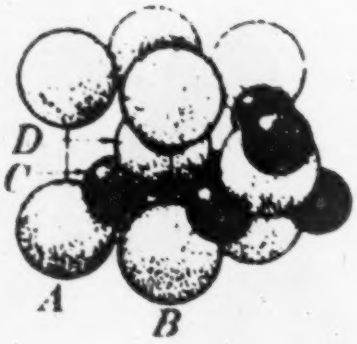
NOT AN EXAMPLE



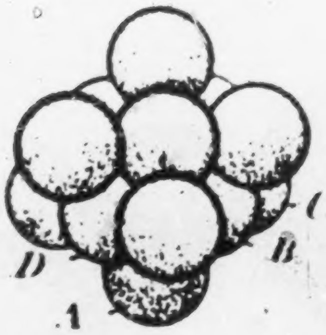
EXAMPLE



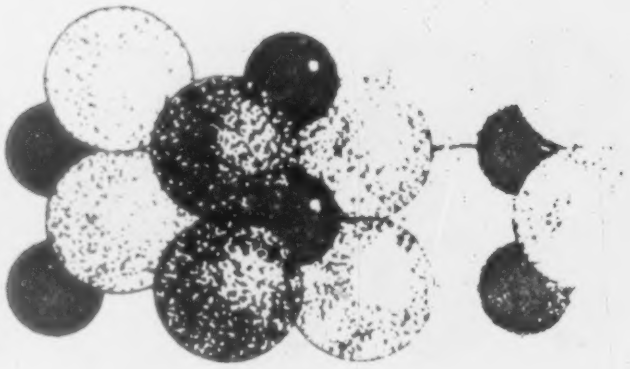
NOT AN EXAMPLE



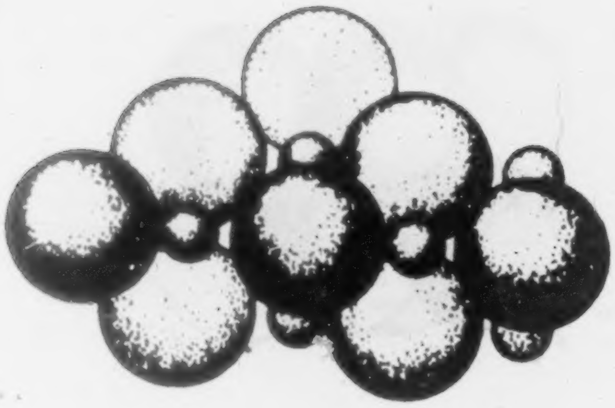
EXAMPLE



NOT AN EXAMPLE

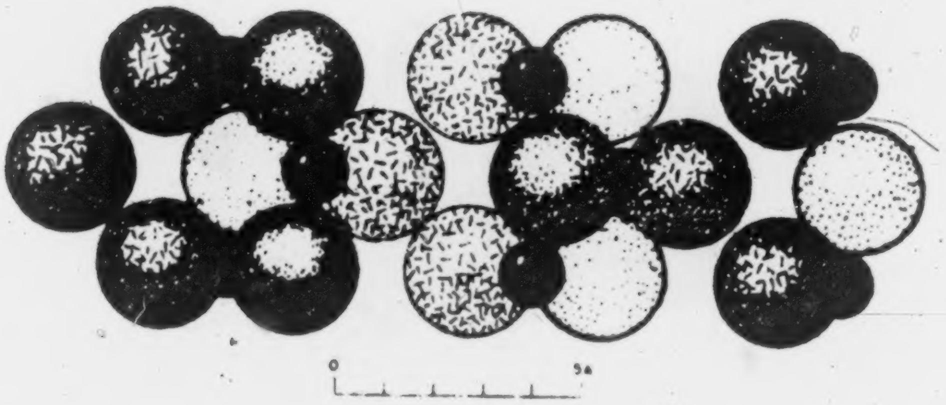


EXAMPLE

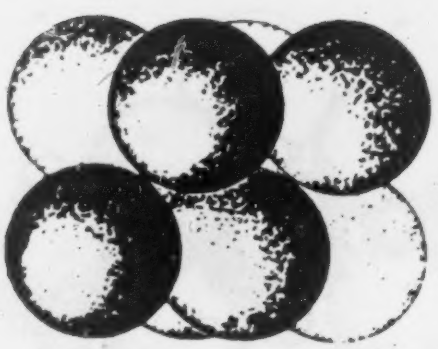


NOT AN EXAMPLE

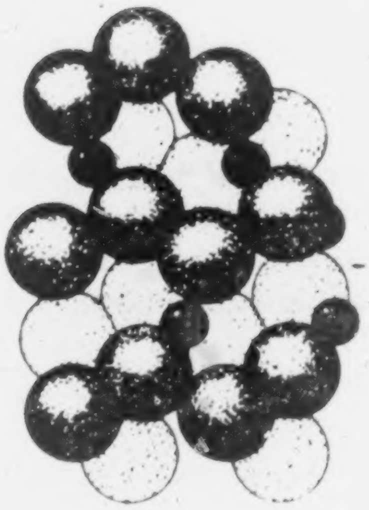
156



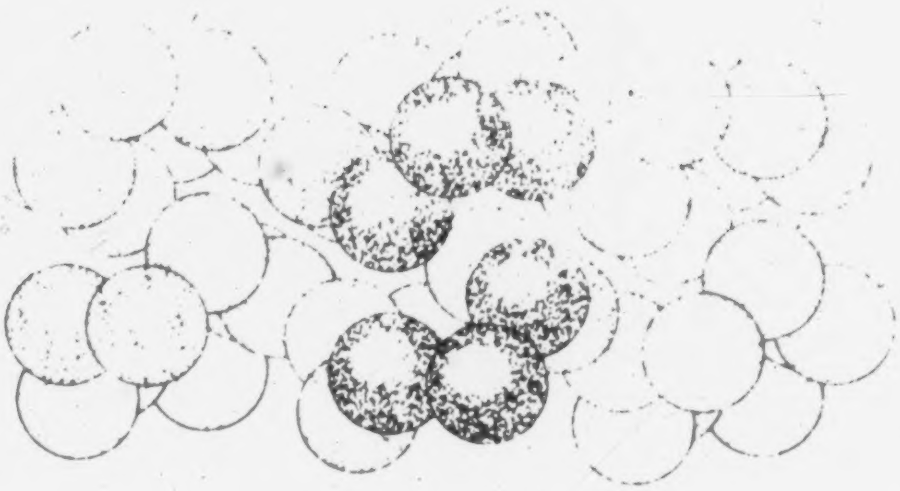
EXAMPLE



NOT AN EXAMPLE



EXAMPLE



NOT AN EXAMPLE

Appendix K Introduction to Nonrelevant Task

Your program will consist of reading selections from an article by Lee J. Cronbach. The article "The Role of the University in Improving Education," presents some very sound information which all students in educational psychology should be aware of.

When you have finished reading the above, please continue to the next page.

Appendix L Nonrelevant Task

The cry today is for instant improvement of the schools, and there is pressure for wholesale dissemination and development activities without the necessary prior research. The educator from the field invariably asks the university to help him with today's problems, and that is understandable. It is understandable, but more deplorable, that Congress spasmodically lashes out with crisis-oriented legislation but is unsympathetic to balanced, across-the-board, long-range plans. It is tragic that in the U.S. Office of Education the Bureau of Research has thrown its forces heavily on the side of "practical products" and dissemination. While the USOE is a passive patron of basic research, it has done nothing to formulate and sell to Congress a policy that will promote the healthy development of basic investigation.

What the USOE views as research is well illustrated by an early 1966 press release in which it reports enthusiastically on the use of electronically compressed speech to teach the blind, the teaching of first-graders by tape recorders, and teaching third-graders to sing medieval plainsong. Some ad hoc novelties such as these have practical value, and they deserve a trial. But the research program ought to have the higher objective of reexamining educational ideas and the underlying of mental development is true, and what does it imply for educators? How can we account for growth in ability to form elaborate sentence structures? How does motivation for achievement develop? What part does personal identification with the teacher play in forming character and interest? And so on.

Massive dissemination encourages faddism in education. What sounds like a good idea is launched nationwide long before it has been determined that the methods used are really suitable. There is no evidence to justify, for example, the California legislation that requires instruction in foreign language in grades six to eight; the assumptions used to justify the requirements are untested and with the law now a fait accompli, no one is about to test them. The energies of the people who might be giving thoughtful attention to language instruction are diverted into a crash program to write curriculum materials and train teachers. The Head Start program is easier to justify, since it reflects a national acceptance of responsibility for the disadvantaged child, and one cannot waste a generation of children while waiting for research. But psychologists do not know whether any of the intervention programs now being installed can produce lasting benefit to intellectual

Continue on Next Page

development; the progrms differ so radically in their assumptions that we can almost be certain that some of them are wrong.

Innovation for innovation's sake is a false value. It crystallizes a practice prematurely and builds up vested interests that discourage hardheaded inquiry and tend to prevent abandonment of the practice when itsglitter wears off.

Now, what recommendations can we make for the proper use of the university as a center for inquiry into education?

First, institutions outside the university should be developed to carry the main burden of demonstration, dissemination, and educational development, The university should, insofar as possible, withdraw from these activities, though it should make the knowledge of its staff available to those who carry them out. Just as aerospace firms and defense laboratories do the developmental work in those engineering fields and the pharmaceutical companies develop products derived from the fundamental studies of the medical school, so educational products will be engineered in institutions resembling Educational Services, Inc., and the Educational Testing Service. These agencies, and the school systems themselves, should do the bulk of in-service training of teachers, though the universities should continue to transmit new ideas to the professional leaders and especially to those who conduct the in-service training.

Second, research should be largely centered in universities, since only the university has the long-range view that permits detached and penetrating inquiry. It is tempting to think of establishing research wings within the development and dissemination institutions, but the hard fact is that in this generation we need to engage every talented researcher as a trainer of researchers and therefore cannot spare him from the university.

Third, the highest priority should be given to recruitment and training of researchers. This calls for breaking down the barriers that now exist between schools of education and other departments. Solid training in one or more of the behavioral, social, and humanistic disciplines is indispensable for thoughtful educational research. Schools of education alone can rarely give that training.

Fourth, we should identify the youngsters who have the greatest

Continue on Next Page

promise as fundamental investigators and should establish the conditions under which they are most likely to become scientific revolutionaries. This means, first of all, financial support for exploratory and unconventional studies, as well as for the neatly canned studies of normal science. It means encouragement of the high-risk activities that do not always pay off, rather than a count-the-publications reward system that locks a man into pedestrian normal science. It means encouraging the man to work on a modest budget that leaves him free to think about his own data.

The improvement of education rests first of all on commitment to the belief that the life of every individual and every nation, and society as a whole, can be lifted to a higher plane of significance through cultivation of the intellect. But improvement will be slight if educational efforts are illuminated by goodheartedness alone. It is a cruel hoax to hail an unsubstantiated method as a cure for an educational deficiency; to adopt it is only to delay the search for underlying causes and for treatments matched to these causes. Intellect begins to play some role in our educational decisions when we test the claims of each new method by assessing its effects in pilot schools. But the intellect takes up its proper duty when it tells us how education and learning proceed, when it tells us why one approach works and another does not, when it identifies the variables that we must adjust to achieve a prescribed effect. The proper mission of the university is to construct, bit by bit, this theory of instruction and of educational systems, while others work on stopgap empirical solutions for educational problems of the moment.

The race between education and catastrophe is not a 60-yard dash, not a matter for spurts of hot-breathed energy. Our generation has a long lap to run. May our pace be strong and our direction sure.

For the person concerned professionally with the improvement of education, this is a time of exhilaration and of despair--exhilaration because our opportunities have expanded to the point where we can almost say that progress is limited only by our capabilities, despair because our capabilities are limiting indeed. Prodigious demands are placed on the school just because it is the one institution under public control that can deliberately cultivate talent and emotional resources--that can on the one hand give individuals the freedom and tools wherewith they carve out a good life, and that can on the other integrate society around principles of opportunity and justice. The nation is ready to back a heroic effort to accomplish these ends, but we professionals do not know enough about learning and instruction to design the desired reform.

Continue on Next Page

While I can only praise in the highest terms the new commitment to education and the enthusiam with which schools are searching for new practices, I am concerned lest the movement may cause the universities, and particularly their schools of education, to neglect their true and unique function. If those whose first calling is the study of education now put off the robe of the scholar and don the armor of the crusader, they will betray the public by leaving the scholar's badly needed work undone.

Effective educational designs, worth careful development and field trial, can emerge only from a deep understanding of learning and motivation.

In the process by which education is improved, we recognize a sequence of activities that starts with basic investigation of the conditions affecting learning, motivation, and instructional effectiveness; carries on through an engineering phase in which practical procedures are designed, tested, and redesigned until they are truly effective, and ends in a marketing phase in which schools are persuaded to adopt the improved methods and teachers are trained to use them. Research, development, dissemination--all three are necessary to keep the educational system moving forward.

Of the three, research is the most difficult to foster. Multiplying research appropriations will not do the job because insightful research requires training and attitudes that are in very short supply. There are few really excellent persons in educational research careers today.

No university can or should fill its faculty entirely with researchers. But any faculty that now contains a modest number of scholars capable of fundamental thought about educational institutions should assign first priority to cultivation of that scholarly capability. With the most careful concentration of existing talent and development of potential talent, we might build up in this country by 1970 a dozen institutions with well-rounded programs of scholarship in education. This will happen, however, only if the better universities hold down their commitments to development and dissemination in order to give first attention to the research mission.

Continue on Next Page

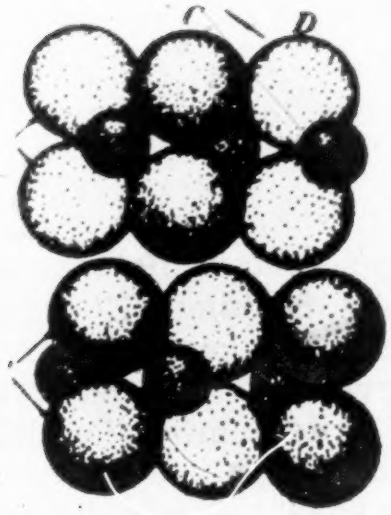
Appendix M Introduction to Posttest

The next few pages contain pictures of crystals. You are to identify each crystal by marking "T" if you think the crystal is an example of RX_2 crystals, or "F" if you think the crystal is not an example of RX_2 crystals. You may spend as much time per item as is necessary to determine the classification. If you did not receive the pretest, mark your answers starting with number 16 on the IBM answer sheet. If you did take the pretest continue marking on the IBM answer sheet with number 16.

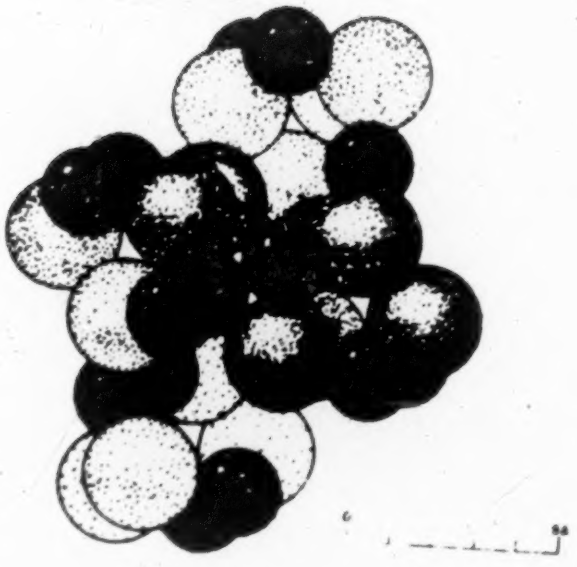
When you have finished reading the above, please continue to the next page.

Appendix N Posttest

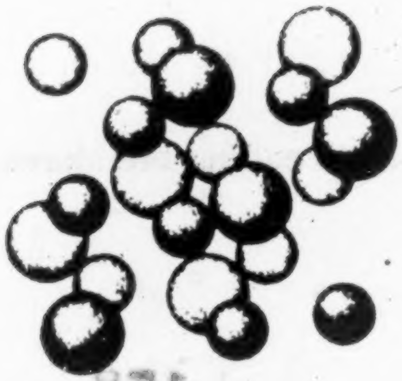
#16



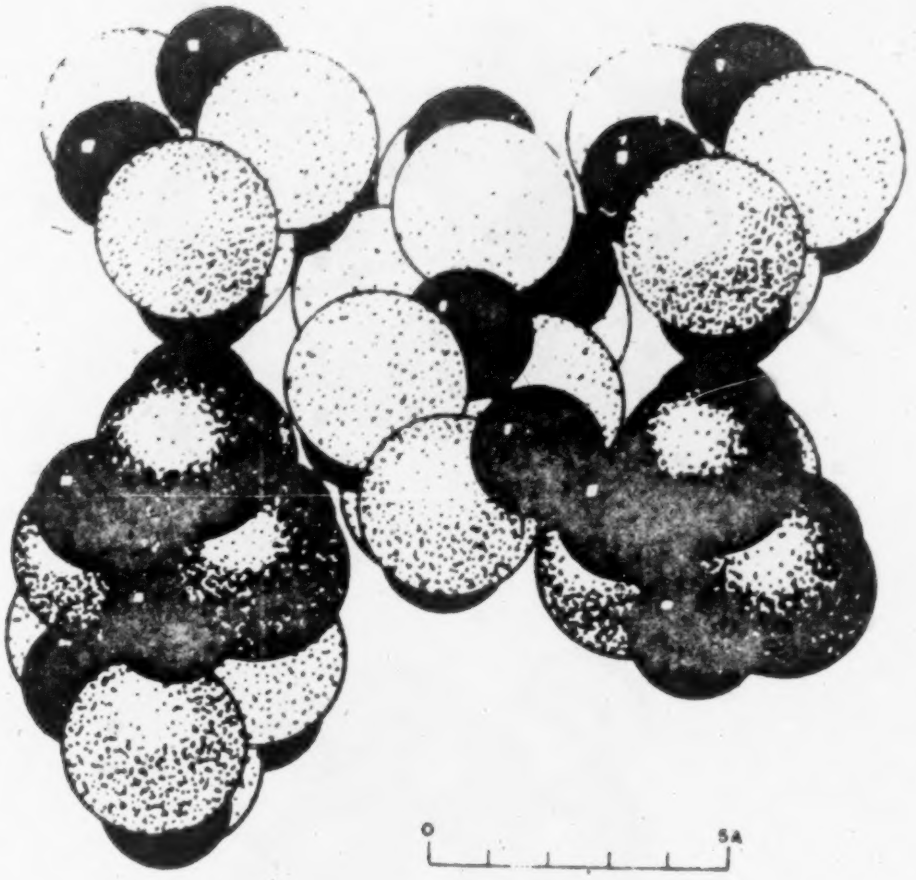
#17



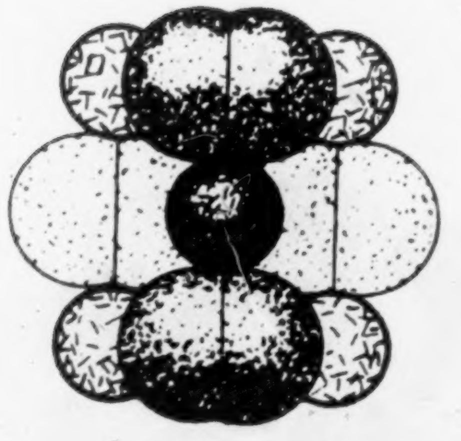
#18



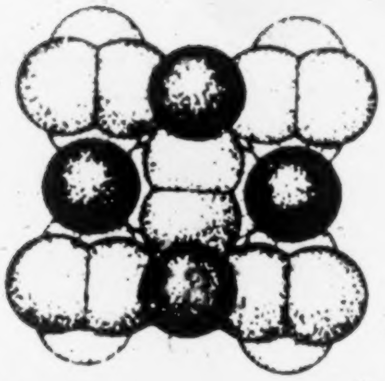
#19



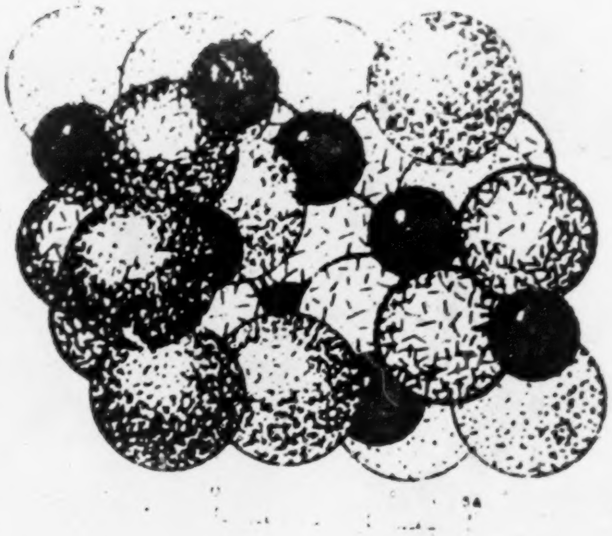
#20



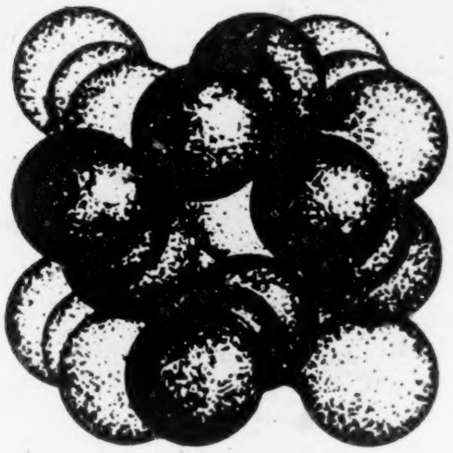
#21



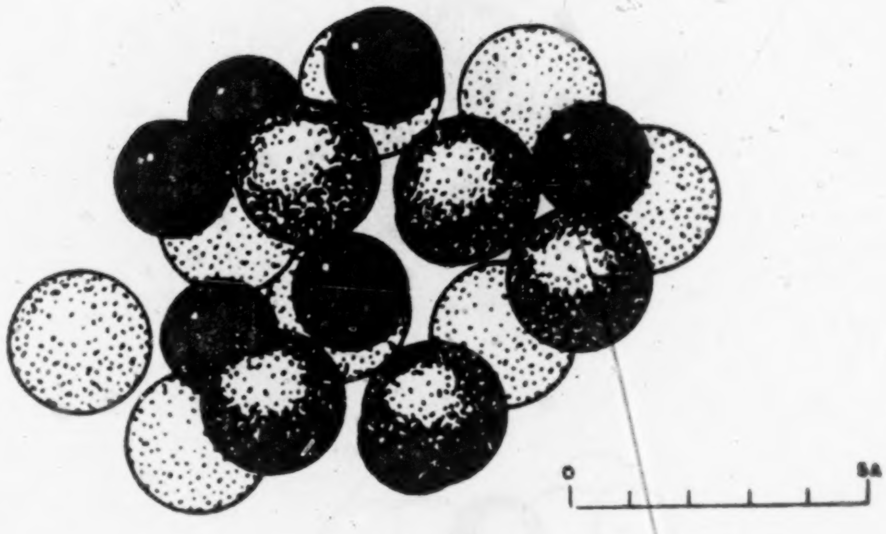
#22



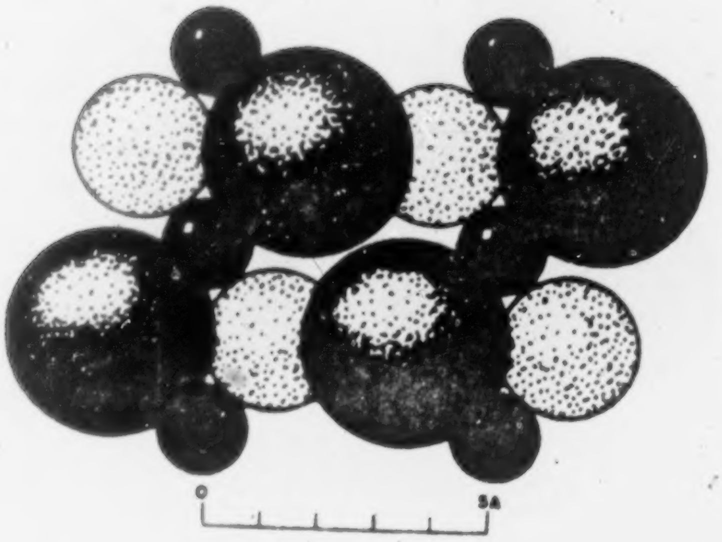
#23



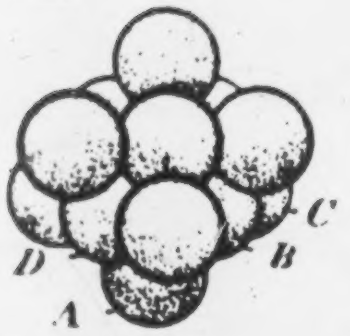
#24



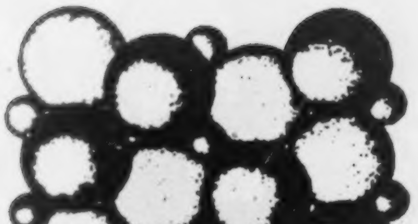
#25



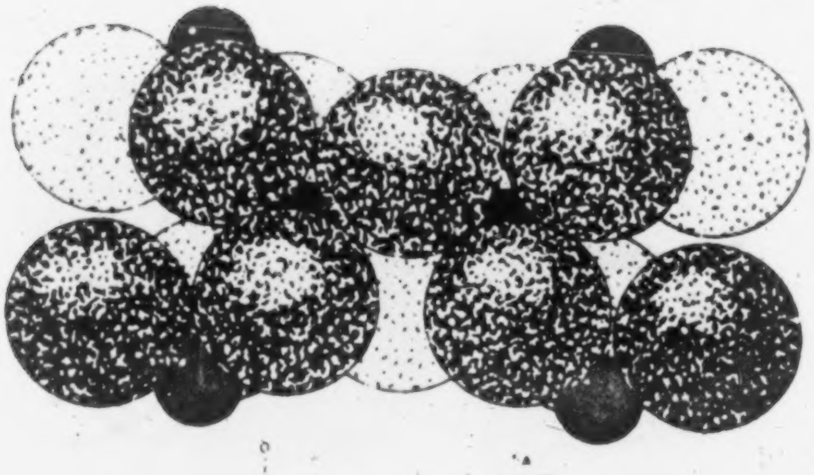
#26



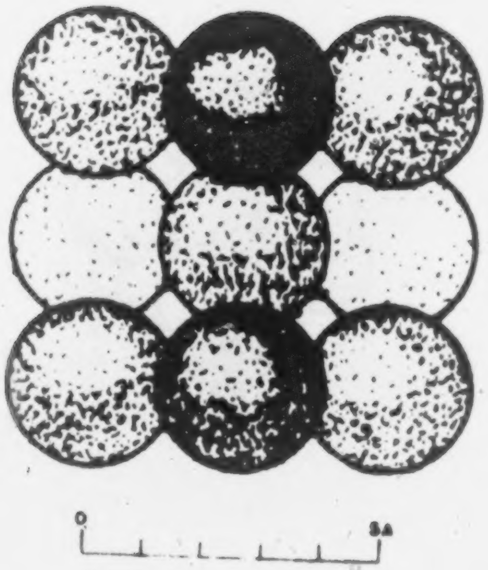
#27



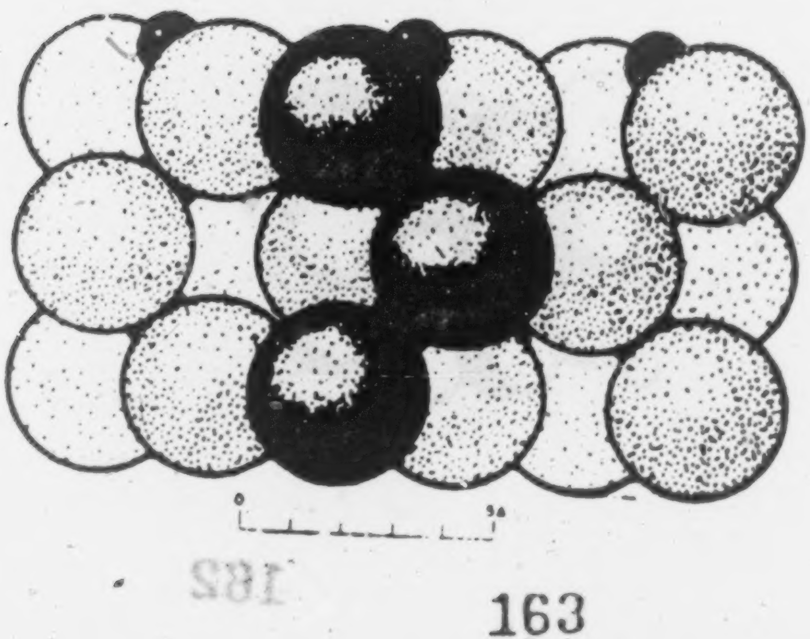
#28



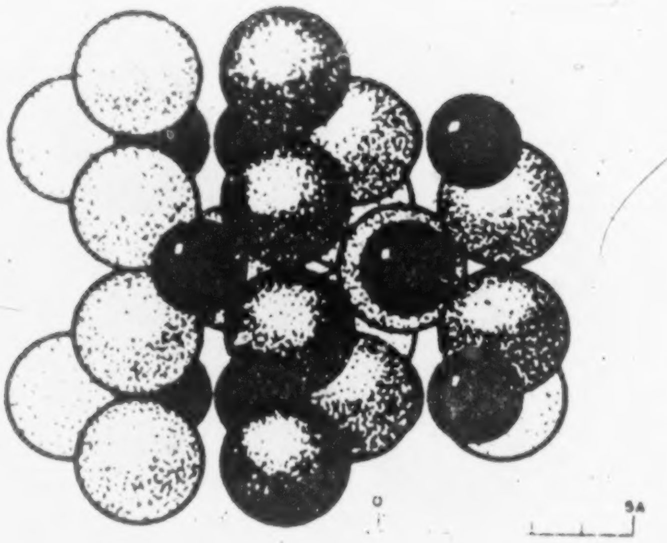
#29



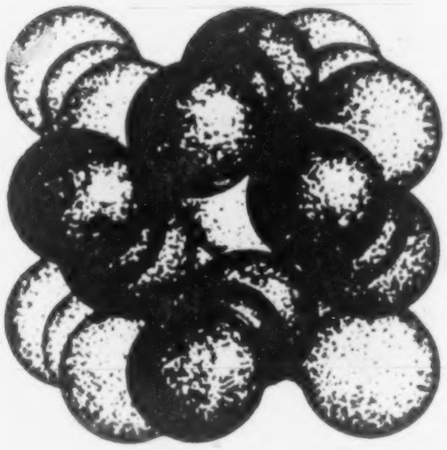
#30



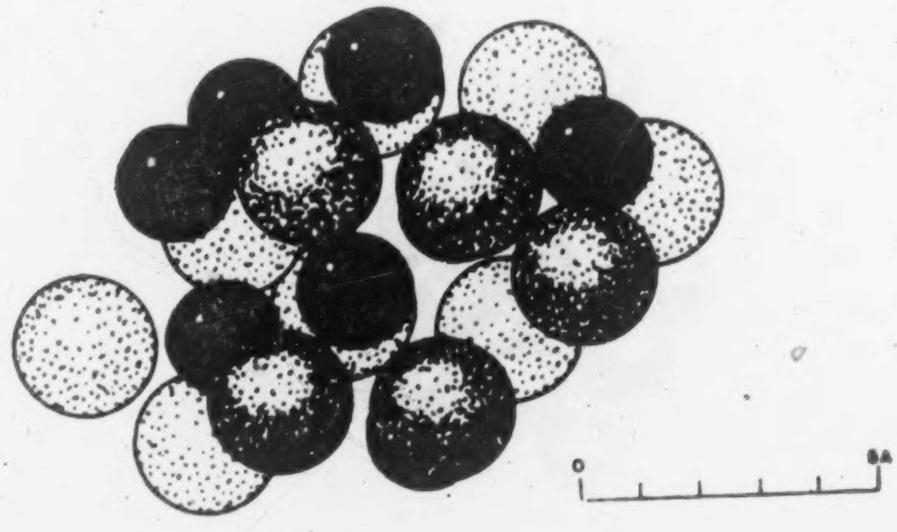
#31



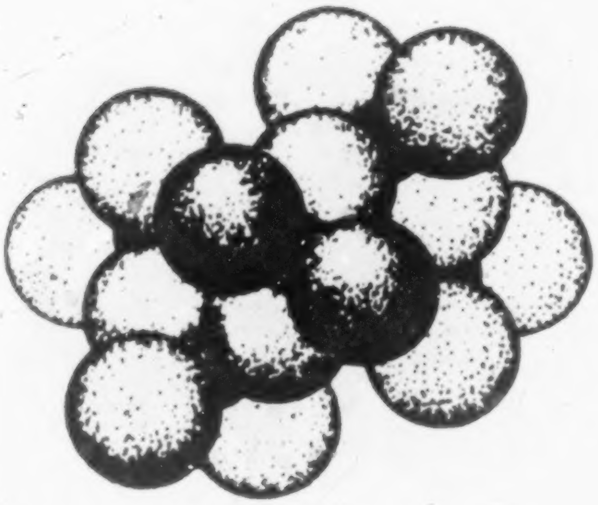
#32



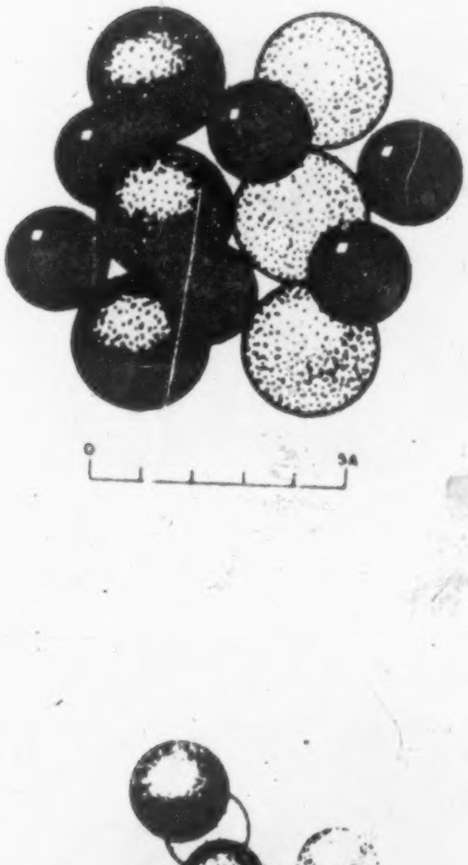
#33



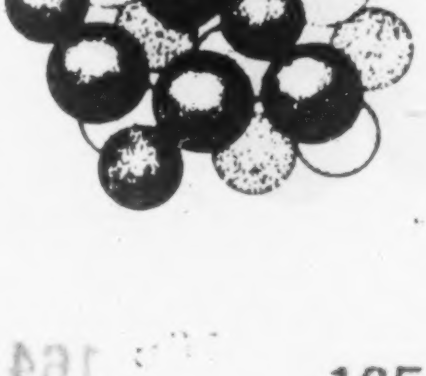
#34

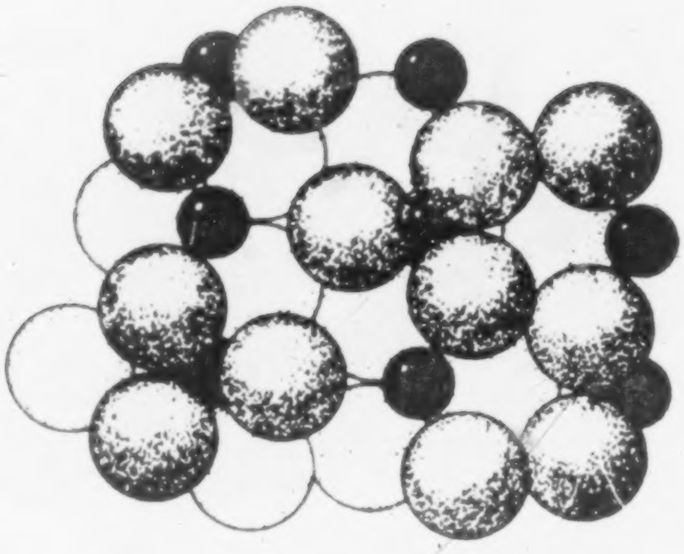


#35

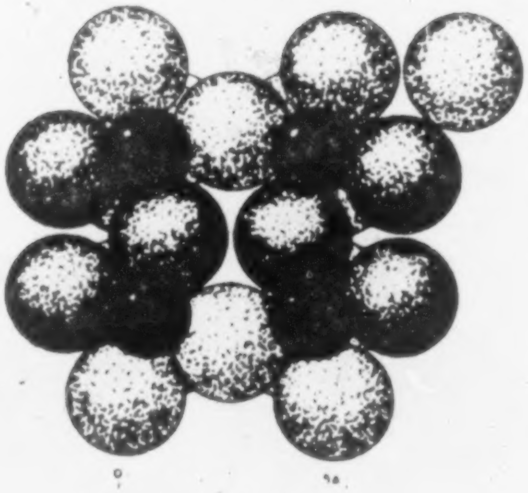


#36

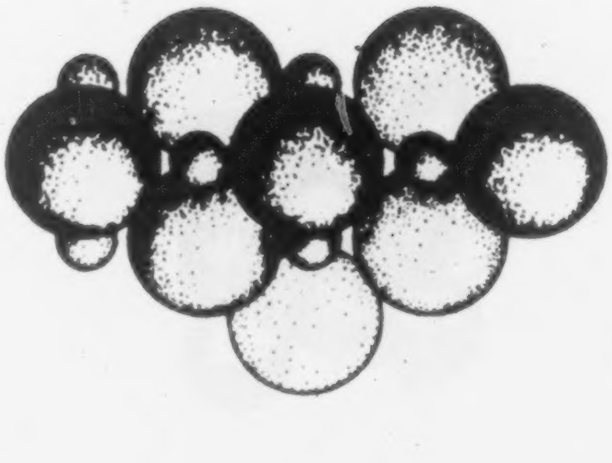




#37

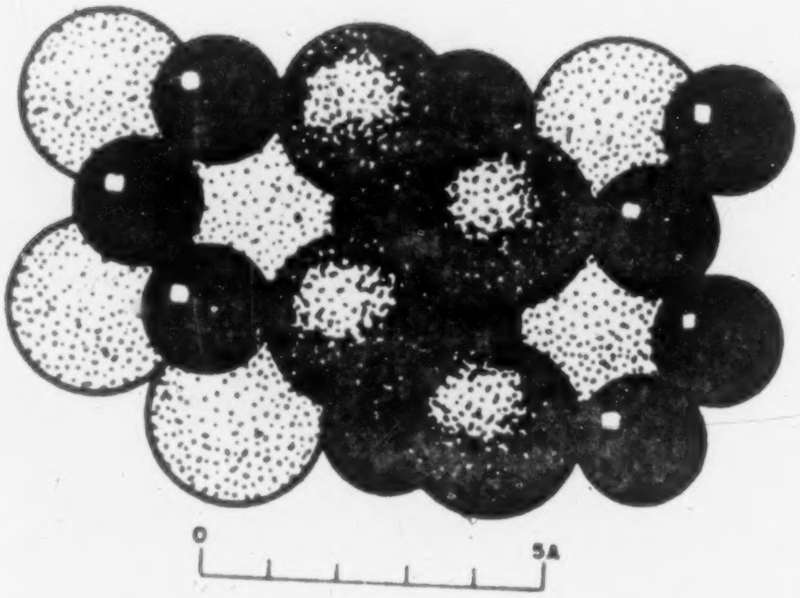


#38

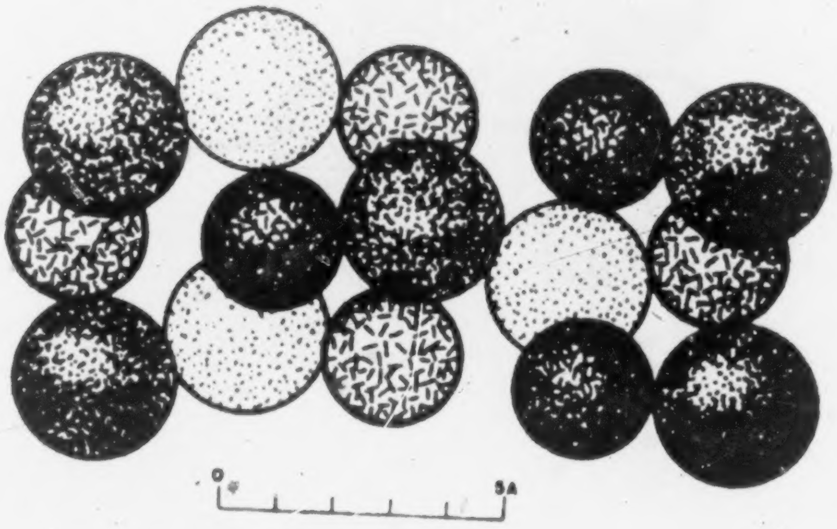


#39

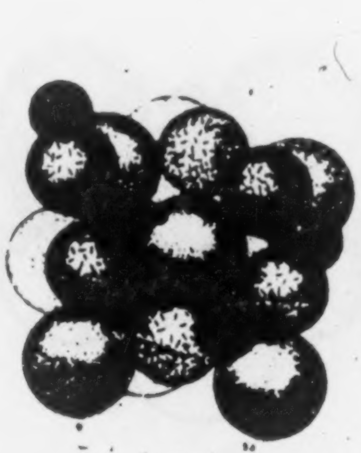
#40



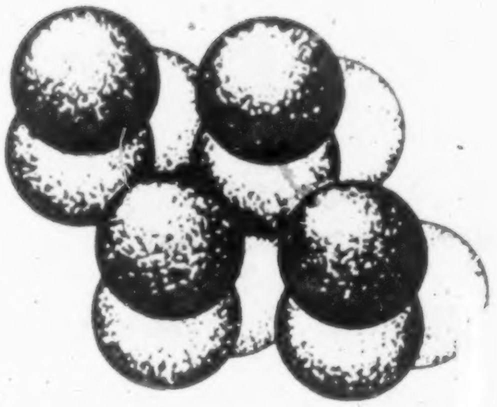
#41



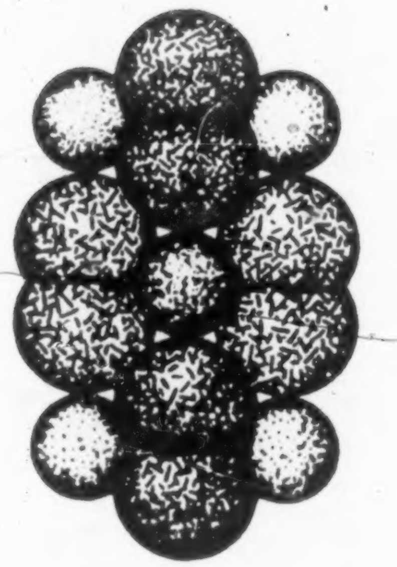
#42



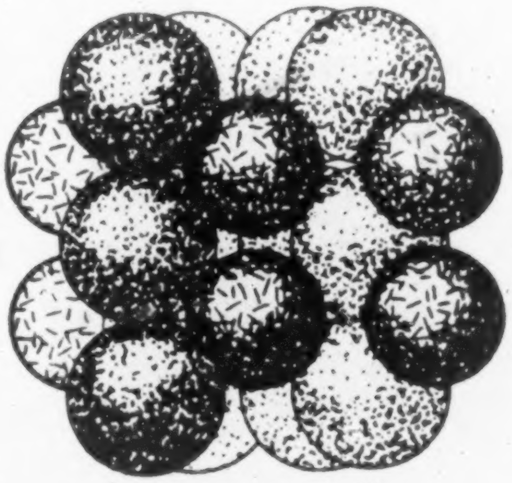
#43



#44



#45



Appendix O
Answers to Posttest

	Overg.	Underg.	Misc.	CC.
16	T	T	T	T
17	T-F	F	T	F
18	T	F	F	F
19	T-F	F	T-F	F
20	T	F	F	F
21	T	F	F	F
22	T	F	T	T
23	T	T	F	T
24	T-F	F	T-F	F
25	T	F	T	T
26	T-F	F	F	F
27	T	T	F	T
28	T-F	F	T-F	F
29	T	T	F	T
30	T	F	T	T
31	T-F	F	T	F
32	T	T	F	T
33	T-F	F	T-F	F
34	T	F	F	T
35	T	F	T	F
36	T	F	F	T
37	T	F	T	F
38	T-F	F	T-F	F
39	T	F	F	F
40	T-F	F	T-F	F
41	T-F	F	F	F
42	T	T	T	T
43	T-F	F	F	F
44	T	F	F	T
45	T-F	F	F	F

Appendix P Posttest Probabilities

Posttest Number	Main book no.	* example	Probability									
			10	20	30	40	50	60	70	80	90	100
16	13	*								81		
17	3				33							
18	39									80		
19	20					42						
20	4						58					
21	14						54					
22	81	*			38							
23	28	*							73			
24	83								74			
25	53	*					58					
26	49								76			
27	36	*						61				
28	54				34	43						
29	21	*				43						
30	35	*			39							
31	12							60				
32	28	*							73			
33	83								75			
34	17	*			35							
35	43					49						
36	40	*			39							
37	59							60				
38	76						56					
39	26				34					84		
40	77						54			84		
41	34					42	54					
42	80	*				42						100
43	8						52					100
44	18	*					52					
45	19							61				

Appendix Q Posttest Construction

Convergent

eg. 13
eg. 81
$\overline{\text{eg}}$. 12
$\overline{\text{eg}}$. 3
$\overline{\text{eg}}$. 20
eg. 28
eg. 17
$\overline{\text{eg}}$. 34
$\overline{\text{eg}}$. 4
$\overline{\text{eg}}$. 39

eg. 64
eg. 35
$\overline{\text{eg}}$. 76
$\overline{\text{eg}}$. 43
$\overline{\text{eg}}$. 77
eg. 21
eg. 18
$\overline{\text{eg}}$. 19
$\overline{\text{eg}}$. 14
$\overline{\text{eg}}$. 49

eg. 80
eg. 53
$\overline{\text{eg}}$. 54
$\overline{\text{eg}}$. 59
$\overline{\text{eg}}$. 83
eg. 36
eg. 40
$\overline{\text{eg}}$. 8
$\overline{\text{eg}}$. 26
$\overline{\text{eg}}$. 11

Appendix R Construction of Tasks

Correct Classification

	Easy	M-Easy	M-Hard	Hard
eg	7*	33	38	70
$\overline{eg}$	10	29	46	20
eg	58	44	41	69
$\overline{eg}$	56	3	78	52

Overgeneralization

	Hard	Hard	Hard	Hard
eg	50	22	41	33
$\overline{eg}$	72	20	46	3
eg	38	58	70	35
$\overline{eg}$	10	42	78	52

Undergeneralization

		Easy	Easy	Easy
eg	25	45	60	58
$\overline{eg}$	20	78	29	52
eg	63	58	13	60
$\overline{eg}$	10	57	61	42

Misconception

eg	1	70	22	32
$\overline{eg}$	10	24	49	42
eg	44	16	19	33
$\overline{eg}$	15	27	26	2

---

*Note--This number is the same as in the main book from the Instance Probability Analysis

Appendix S Conclusion to Program

You are now finished. Thank you very much for your participation in this study. The results should be available sometime this semester. If you are interested, please feel free to contact Dr. David Merrill (ext. 2635), and the information will be given to you.

Please return this booklet to the experimenter and leave quietly.

Appendix T

Analysis of Variance Tables

Correct Classification
Analysis of Variance

	Sum of Squares	Df	Mean Square	F Ratio
Between Groups	263.5600	4	65.8900	4.6867*
Within Groups	1335.6000	95	14.0589	
Total	1599.1600	99		

*p<.01.

Misconception

Analysis of Variance

	Sum of Squares	Df	Mean Square	F Ratio
Between Groups	553.2600	4	138.3150	7.1791*
Within Groups	1830.3000	95	19.2663	
Total	2383.5600	99		

*$p < .01$.

Undergeneralization

Analysis of Variance

	Sum of Squares	Df	Mean Square	F Ratio
Between Groups	290.0400	4	72.5100	3.2991*
Within Groups	2088.0000	95	21.9789	
Total	2378.0400	99		

*$p < .025$.

Overgeneralization

Analysis of Variance

	Sum of Squares	Df	Mean Square	F Ratio
Between Groups	321.9600	4	80.4900	3.3873*
Within Groups	2257.4000	95	23.7621	
Total	2579.3600	99		

*$p < .025$.